# Nonparametric Modelling and Estimation of Stochastic Volatility

Von der Carl-Friedrich-Gauss-Fakultät für Mathematik und Informatik der Technischen Universität Braunschweig genehmigte Dissertation zur Erlangung des Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)

**Andreas Dürkes**

Präsident der Technischen Universität Braunschweig:
Prof. Dr.-Ing. Dr. h. c. Jürgen Hesselbach

Dekan der Carl-Friedrich-Gauss-Fakultät für Mathematik und Informatik:
Prof. Dr.-Ing. Lars Wolf

Gutachter:
1. Prof. Dr. Jens-Peter Kreiß
2. Prof. Dr. Michael H. Neumann

eingereicht am: 1. November 2006
Tag der mündlichen Prüfung: 14. Dezember 2006

To my father

# Contents

# Chapter 1

# Introduction

The aim of financial mathematics and financial time series analysis is to describe the behaviour of financial markets. Financial data consist of a time series of prices of a certain asset which are e.g. evaluated daily, hourly, minutely or even tick-by-tick, which means more than secondly, i. e. almost infinitely often.
Prices of such derivatives are often transformed to returns or log-returns, namely

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

or

$$R_t = \log(P_t) - \log(P_{t-1}),$$

respectively. A Taylor series expansion yields, that returns and log-returns are approximately the same, if the relativ price changes are small.
A well known and often used model for option-prizing and to describe financial data is the so called Black-Scholes-Model, also known as Samuelson- or Samuelson-Black-Scholes-Model. Samuelson (1973) [86] and Black and Scholes (1973) [17] modelled financial data as a geometric Brownian motion

$$P_t = P_0 \exp\left(\sigma W_t + \mu t - \frac{\sigma^2}{2} t\right), \quad t \in I\!R. \tag{1.1}$$

This model is also written in the form

$$dx(t) = (\mu + \beta\sigma^2)dt + \sigma dW_t, \tag{1.2}$$

where $x(t)$ would model the logarithm of an asset price and $\beta$ denotes a generalised risk-premium.
Discretizing this model and changing to log-returns leads to the following model in discrete time:

$$R_t = \mu + \beta\sigma^2 + \sigma\epsilon_t, \tag{1.3}$$

with $\epsilon_t = W_t - W_{t-1} \sim \mathcal{N}(0,1)$.
The Brownian motion $W_t$ is a stochastic process, which describes movements of

small particles in a liquid and is regarded as having been discovered by Robert Brown (1827) [23]. In 1880 Thorvald Nicolai Thiele was the first one to describe this process mathematically by describing the residuals arising from least square estimators [101]. The first one to introduce the Brownian Motion to financial mathematics was Louis Bachelier (1900) in his Ph.D. thesis "Théorie de la spéculation" [8]. Maybe his idea of modelling stock-prices as Brownian motions with drift was too far ahead of its time, because he only received a lower grade, cf. Courtault, Kabanov, Bru, Crepel (2000) [29] and Taqqu (2001) [96].
It was Albert Einstein (probably not knowing the work of Bachelier) in 1905 [38] who attracted attention to this stochastic process and defined it in the contemporary way. A collection of Einstein's work on the Brownian Motion (including [38]) can be found in the small booklet "Investigations on the Theory of the Brownian Movement" [39].

Already in the early second half of the last century empirical studies of Mandelbrot (1963) [68] and Fama (1965) [43] have shown, that the assumption of constant volatility - as made in the equations (1.1), (1.2) and (1.3) - can not be maintained. They deduced, that there is a certain dependence structure among the data, that volatility changes in time and that the data are heavy-tailed.
It was already clear to Black and Scholes, that homogeneity was a unrealistic assumption. They wrote in 1972, cf. [16]: "... there is evidence of non stationarity in the variance." In so called stochastic volatility models equation (1.2) is thus generalised to allow the volatility term to be stochastic and to vary over time. This generalised model reads as follows:

$$dx(t) = (\mu + \beta\sigma^2(t))dt + \sigma(t)dW_t, \tag{1.4}$$

where $\sigma^2(t)$, which is usually called instantenous or spot volatility, is assumed to have locally square integrable sample paths, while being stationary and stochastically independent of the Brownian motion $W(t)$. By allowing the spot volatility to be random and serially dependent these models overcome a major failing in the Black-Scholes option pricing approach, cf. Hull and White (1987, 1988) [60, 61] and Heston (1993) [57], but the assumption, that the driving process is Gaussian is still maintained. Among others both articles of Hull and White are contained in the book "Hull-White on Derivatives" [62], where various articles of Hull and White on derivatives are collected.
A comprehensive overview of how to model financial time series and the so called stochastic volatility models can be found in the work of Taylor (1986, 1994) [98, 99], Ghysels, Harvey and Renault (1996) [52] and Shephard (1996) [87], a discussion on publications and reprints of collected articles published in this context are given in Shephard (2005) [88]. Statistical aspects of these stochastic volatility models are e.g. studied in Barndorff-Nielsen and Shephard (2001, 2002) [10, 11].

A model class to describe discrete financial data, which does not need the assumption of normally distributed innovations has been proposed by Robert Engle in 1982, cf. [41]. This model class, he received the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel 2003 for, was called "Autoregressive Conditional Heteroscedastic" (ARCH(p)) and is written as follows:

$$\begin{aligned} X_t &= \sigma_t e_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-1}^2, \quad t \in I\!R, \alpha_0, \ldots, \alpha_{p-1} \geq 0, \alpha_p > 0, \end{aligned}$$

with $e_t$ centered and i.i.d., but as already mentioned not necessarily normally distributed.
This model has been generalised by Tim Bollerslev in 1986, who modelled the actual variance as a weighted sum of the past returns $X$ and variances $\sigma^2$, i.e.

$$\begin{aligned} X_t &= \sigma_t e_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-1}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-1}^2, \\ & \qquad t \in I\!R, \alpha_0, \ldots, \alpha_{p-1}, \beta_1, \ldots \beta_{q-1} \geq 0, \alpha_p, \beta_q > 0, \end{aligned}$$

and called his model "Generalized Autoregressive Conditional Heteroscedastic" (GARCH(p,q)) [18].
A shortcoming of standard ARCH and GARCH models is, that they can not describe an asymmetry of volatility, since $\sigma_t^2$ is just a function of the squared past, i.e. they do not imply the so called leverage effect, which says, that negative returns lead to a stronger tendency for higher future volatility than positive returns do, cf. Black (1976) [15] and Nelson (1991) [78].
There are plenty of generalisations of ARCH and GARCH. A survey of this model-class is e.g. given in Bollerslev, Chou and Kroner (1992) [19] and Bollerslev, Engle and Nelson (1986) [20]. A comprehensive collection of reprints of some articles in this context has been edited by Robert Engle in 1995, cf. [42].

Throughout this work we will consider returns (or log-returns) of financial data, which can be modelled as uncorrelated random variables, whose conditional variance $\sigma_t$ is realised by a non observable (hidden) stochastic process

$$R_t = \mu + \sigma_t V_t. \tag{1.5}$$

Although we will assume, that the parameter $\mu$ - the expectation of the returns $R_t$ - is known. From applications it can be justified that this expectation is close to zero.
Due to the above mentioned shortcomings of standard ARCH and GARCH models we will assume that the model structure of the hidden volatility process

is nonparametric. Furthermore we will assume, that the distribution of the innovations $V_t$ is unknown. Thus the considered models generalise on one hand parametric autoregressive random variance models, such as the various ARCH- and GARCH-Models, which quite successfully have been applied to financial time series, and on the other hand nonparametric stochastic volatility models for which the distribution of the innovations of the returns is assumed to be known, as in the discrete case e.g. been considered in the work of Franke, Härdle and Kreiss (2003) [48] and in van Es, Spreij and van Zanten (2005) [40].

Throughout the chapters 2 - 4 we will consider a discrete return- or log-return-process, which can be described by an equation of the type of (1.5). The following nonparametric structure of the volatility process $\log \sigma_t$, which throughout the whole work is denoted by $\xi_t$, is assumed:

$$\xi_t = m(\xi_{t-1}) + \eta_t, \tag{1.6}$$

with independent and identically distributed (i.i.d.) innovations $(\eta_t)_{t \in \mathbb{N}}$.

The aim of this work is, to give an estimator of the density of the non-observable volatility and of the regression function $m$. If the volatilities $\xi$ could be observed, one would estimate the stationary density of $\xi$ by usual kernel estimators and the regression-function by a so-called Nadaraya-Watson estimator as introduced independently by Nadaraya and Watson in 1964 [77, 104]. This situation has e.g. been considered by Robinson in 1983 [85] in the univariate and by Masry in 1991 [69] in the multivariate case. An account of the various techniques, that have been proposed to estimate an unknown density can be found in Silverman (1986) [89]. A survey on kernel estimation including kernel regression is given in Wand and Jones (1995) [102], a quite comprehensive overview about "All of Nonparametric Statistics" including kernel density estimation and kernel regression in Wasserman (2006) [103].

Unfortunately, as already mentioned above, we can only observe the volatility disturbed by a noise term, i.e. by centering and taking logarithms of the returns we observe a disturbed random variable $X_t$ of $\xi_t$, given by

$$X_t = \xi_t + \epsilon_t. \tag{1.7}$$

In statistics one is often confronted to the situation not to be able to observe the values one is interested in. One of the eldest works known to the author which mentions this fact is a short article of A. S. Eddington (1913) [36] dealing on astronomical investigations. Usually if you are confronted to this situation in the regression problem, it is assumed to know the distribution of the errors. Then the distribution of $\xi_t$ can be estimated via a division of the empirical characteristic function by the Fourier transform of the error density, a so-called deconvolution kernel estimate. Various works dealing on this context are Carroll and Hall (1988) [26], Liu and Taylor (1989) [66], Devroye (1989) [32] Stefanski and Carroll (1990, 1991) [92, 93], Stefanski (1990) [91], Zhang (1990) [105], Fan (1991, 1991, 1993) [44, 45, 46], Fan and Truong (1993) [47] and Masry (1991, 1993, 1993) [70, 71, 72]. A different approach has already been given in Mendelsohn and Rice (1982) [76], where the disturbed density has been estimated and the density of interest has been determined by a minimizing problem.

In our situation knowing the error distribution would mean to assume to know the distribution of the innovations $V_t$. Usually the innovations are supposed to be normally distributed. Postulating this normality, it can be shown, that the usual deconvolution estimators known from the regression problem can successfully be applied to the related time series situation, cf. Franke, Härdle, Kreiss (2003) [48]. It should be mentioned that for this approach the assumption of normally distributed innovations $V_t$ of the return process is not necessary, but it definitely is necessary to assume that the distribution is completely known, which is of course rather unrealistic in real applications.
Since - as already highlighted above - log-returns possess a leptokurtic distribution with much fatter tails than normal distributions, especially the normal distribution seems to be questionable for describing real financial data situations. This fact has already been discussed in many early publications (cf. Praetz (1972)[83], Blattberg and Gonedes (1974)[13] or Taylor and Kingsman (1979)[100]). The leptokurtic distribution of log-returns is emphasized in RiskMetrics-Technical Document [84] as well, even though RiskMetrics uses the normality assumption.

One aim of this work is to dispense with the assumption of knowing the distribution of the innovations $V_t$ of the return process in order to construct consistent estimators of the underlying nonparametric structure of the volatility process.
The problem of not knowing the distribution of the observation-errors has been considered by Diggle and Hall in 1993 [33]. They suppose, that information of the error distribution can be drawn from an additional experiment and propose to use the standard kernel deconvolution technique with the empirical characteristic function of the errors inserted for their unknown characteristic function. The effect of estimating the error density on rates of convergence has been studied by Neumann in 1997 [79]. More work on this topic has been done by Efromovich in 1997 [37] and Meister in 2004 [74].
Cases, in which the error density and the distribution of interest are of different characteristics, have been considered by Meister in 2004 [75] and Butucea and Matias in 2005 [25]. Both identify the true error function from the tail behaviour of the characteristic functions of the observations.

It seems to be quite well accepted to assume, that the volatility process $\sigma_t$ varies much more slowly than does the process $R_t$ or $V_t$. This fact has already been

mentioned in the paper of Taylor (1982) [97], in which the daily sugar prices for the period 1961 - 1979 are considered. Another hint towards this fact is given in the RiskMetrics-Technical document [84], where the USD/DEM returns are considered in order to demonstrate, that volatility clustering occurs, i.e. that periods of high volatility are clearly separated from periods with lower volatility. To attain our aim to dispense with the assumption of knowing the error-distribution as well, we will make use of this awareness of slower movement of the volatility process. In chapter 2 we consider - as an extreme case - a situation in which we are able to observe a short panel of return data which rely on exactly the same volatility. It will be seen that at least two returns with identical volatility are needed to carry through our estimation procedure, which follows ideas of Horowitz and Markatou (1996) [59]. Furthermore we assume for this model, that the characteristic function of the innovations of the volatility process is strictly positive. For example this assumption is fulfilled, if the innovations can be written as a sum of two independent symmetric random variables.
Chapter 3 considers a situation in which the observation errors converge to zero in probability with increasing sample size. In such a situation we get completely rid of the deconvolution dilemma, because we are now in the situation of an errors-in-variables model, cf. Zwanzig (1999, 2003), [108, 109]. Usual kernel smoothing methods become applicable in this case.
In chapter 4 we will present two model-assumptions in which the conditions of chapter 3 are fulfilled. In both cases we assume that we can observe the asset of interest almost infinitely often and that with increasing sample-size we take more observations. In section 4.1 we consider intra-day log-returns, which are defined via direct neighbour observations, and assume that by using this intra-day returns we can estimate the variance of the day. There we are in a situation comparable to so-called realised and integrated variances in continous models. In section 4.2 we consider inter-day returns and assume that there exists a daily-mean-volatility, which follows the autoregressive structure and can be estimated by taking an increasing number of observations.

Throughout the chapters 2 - 4 we will suppose, that the following general assumptions on the volatility process and the autoregression (1.6) are fulfilled:

**1.1 Assumptions.**

**A1** $\limsup_{|x|\to\infty} \left|\frac{m(x)}{x}\right| < 1$,

**A2** $f_\eta$, *the density function of* $\eta$, *which is assumed to exist, is strictly positive on all compact sets.*

**A3** $\mathrm{E}(\eta_t) = 0$ *and* $\mathrm{E}(\eta_t^2) < \infty \; \forall\, t \in I\!N$.

A1 and A2 ensure that a strictly stationary solution of (1.6) exists. Since we

assume stationarity for the underlying volatility process, A1 and A2 also ensure that $(\xi_i)_{i\in\{1,\dots,T\}}$ is geometrically ergodic (cf. Doukhan (1994) [34], page 107 Proposition 6), which implies geometrical $\beta$- and $\alpha$-mixing if the chain is stationary. For the definition of $\alpha-$ and $\beta-$mixing compare Doukhan (1994) [34] or Bosq (1996) [21].
We denote by $\alpha_\xi(k)$ the $\alpha$-mixing coefficient of the process $(\xi_t)_{t\in\mathbb{N}}$ and by $\rho_\xi$ a positive real number less than 1 with

$$\alpha_\xi(k) \leq \rho_\xi^k. \tag{1.8}$$

The assumption of a nonvanishing density (cf. A2) of the innovations can be relaxed - details can be found in Franke, Kreiss, Mammen and Neumann (2002) [50].
In order to apply the results of chapter 3 estimators of the volatility were simulated and the introduced techniques were used to estimate the stationary density and the autoregression function of the volatility. These estimations - always confronted with the real autoregression function - are given in chapter 5. The techniques were also applied to real data. Some estimates taken from DAX-Data of 1997 based on the models of chapter 4 are presented in chapter 6.
In chapter 7 we will consider a nonparametric GARCH(1,1) model, which does not comprehend the above mentioned symmetry-dilemma. We will show the existence of a process following this structure, give an estimator of the stationary distribution of such a process and show the asymptotic normality of this estimator. Lastly we will sketch how the nonparametric regression function could be estimated in this model.
Finally it should be mentioned, that all proofs of chapter 2, 3 and 7 are deferred to appendixes A, B and C, respectively.

**Acknowledgement**
I would like to take advantage to express my gratitude to Prof. Dr. J.-P. Kreiss for proposing this interest subject and supporting the present work in the most valuable way.

# Chapter 2

# Multiple Returns based on the same Volatility

In this chapter we consider in detail the following stochastic volatility model for (e.g. daily) returns $R_t$, $t = 0, 1, 2, \dots$ of some financial process in discrete time:

$$R_t = \mu + \sigma_t V_t, \tag{2.1}$$

with i.i.d. random variables $(V_t)_{t \in \mathbb{N}}$, a constant mean $\mu$ and a stochastic volatility $\sigma_t$. Furthermore it is assumed that $\xi_t = \log \sigma_t$ follows the first order nonparametric autoregressive model (1.6).
As argued in the introduction, we assume, that volatility varies much more slowly in time than does the return process itself. As an extreme situation we assume here, that we are able to observe a short panel of returns on the same day, which rely on exactly the same volatility, i.e. we suppose that we are able to observe for each $j = 1, 2, \dots$

$$R_{j,m} = \mu + \sigma_j V_{j,m}, \quad m = 1, \dots, M, \tag{2.2}$$

where the subscript $m$ indicates that we have the $m$-th observation based on the actual (the $j$-th) volatility. Of course this is a rather restrictive assumption, but the intention behind this assumption is to see, how far we can improve on the situation $M = 1$, in which we have to assume that we completely know the distribution of $V_t$ in order to receive a consistent estimate of the function $m$, which determines the stochastic behaviour of the volatility process $(\xi_t)_{t \in \mathbb{N}}$. We will see in the following that a small number of at least $M = 2$ contiguous returns with exactly the same volatility allows us to drop the assumption of knowing the distribution of the errors $V_t$.
Defining

$$X_{j,m} := \frac{1}{2} \log(R_{j,m} - \mu)^2 \quad \text{and} \quad \epsilon_{j,m} := \frac{1}{2} \log V_{j,m}^2, \tag{2.3}$$

we come to the following model

$$X_{j,m} = \xi_j + \epsilon_{j,m}, \quad j = 1, 2, \dots, \quad m = 1, \dots, M \tag{2.4}$$

and

$$\xi_j = m(\xi_{j-1}) + \eta_j, \quad j = 1, 2, \ldots \tag{2.5}$$

It is possible to assume, that the variables $\epsilon_{j,m}$ have expectation zero, since changing the $\epsilon_{j,m}$ by an additive constant means changing $V_{j,m}$ by a multiplicative constant, which is possible because we anyway can separate $V_{j,m}$ and $\sigma_j$ only up to a multiplicative factor.
The usual nonparametric deconvolution estimator for this situation, cf. Franke, Härdle, Kreiss (2003) [48], essentially needs to know the distribution of the errors $\epsilon_{j,m}$ in (2.4) in form of the characteristic function $\phi_\epsilon$.
The estimator of $m$ investigated in Franke et al. (2003) [48] reads as follows

$$\hat{m}(x) = \frac{\frac{1}{TM\lambda_T} \sum_{j=1}^{T} \sum_{m=1}^{M} K_T\left(\frac{x - X_{j,m}}{\lambda_T}\right) X_{j+1,m}}{\hat{f}_\xi(x)}, \tag{2.6}$$

where

$$\hat{f}_\xi(x) = \frac{1}{TM\lambda_T} \sum_{j=1}^{T} \sum_{m=1}^{M} K_T\left(\frac{x - X_{j,m}}{\lambda_T}\right) \tag{2.7}$$

denotes an estimator of the stationary density of the autoregressive process $(\xi_j)_{j \in I\!N}$. $K_T$ denotes the following so-called deconvolution kernel

$$K_T(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\tau x} \frac{\phi_U(\tau)}{\phi_\epsilon(\tau/\lambda_T)} d\tau, \tag{2.8}$$

where the characteristic function $\phi_U$ is defined in A7 below. Notice, that $\hat{f}_\xi(x)$ can also be written in the following form

$$\hat{f}_\xi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\tau x} \phi_U(\lambda_T \tau) \frac{\hat{\phi}_X^T(\tau)}{\phi_\epsilon(\tau)} \, d\tau, \tag{2.9}$$

with

$$\hat{\phi}_X^T(\tau) := (T \cdot M)^{-1} \sum_{j=1}^{T} \sum_{m=1}^{M} \exp(i\tau X_{j,m}). \tag{2.10}$$

Because of the special panel structure of our data we can refer to Horowitz and Markatou (1996) [59] or Horowitz (1998) [58] and consider for $m \geq 2$

$$Z_{j,m} := X_{j,m} - X_{j,1} = \epsilon_{j,m} - \epsilon_{j,1} \,. \tag{2.11}$$

Since we assume the independence of $\epsilon_{j,m}$ and $\epsilon_{j,1}$ (for $m \geq 2$) the characteristic function of $Z_{j,m}$ (which are observable random variables) satisfies

$$\phi_Z = |\phi_\epsilon|^2 \,. \tag{2.12}$$

Thus we are able to estimate the absolute value of $\phi_\epsilon$ by the square-root of the empirical characteristic function of the $Z_{j,m}$. In the case of $\phi_\epsilon$ being real and non-negative this directly leads to an estimator of $\phi_\epsilon$, in which we are interested. In the general case of a not real-valued characteristic function, Horowitz and Markatou sketched an idea of how to estimate the phase-function, if $M > 3$. Under the assumption, that $\phi_\xi$ and $\phi_\epsilon$ do not vanish, Li and Voung (1998) [65] showed, that they can be identified up to a location shift from the joint characteristic function of the random variables $X_{j,m}$. Their result is essentially based on Kotlarski (1962) [63]. Recent work in this context can also be found in Neumann (2006) [80], where the pair of characteristic functions is fitted by a minimum distance method.
Replacing $\phi_\epsilon$ in (2.8) and (2.9) by a thoroughly modified version of the estimator given in (2.12) we indeed obtain a consistent estimator of the autoregression function $m$ of the hidden stochastic volatility process as we will see in the following. To carry through the technical proof of such a result we need the following additional assumptions.

**2.1 Assumptions.**

**A4** *$m$ is twice continuously differentiable with*

$$\left\|m^{(k)}\right\|_\infty < \infty, \qquad k = 1, 2.$$

*Notice that $m$ itself is not necessarily bounded.*

**A5** *$\phi_\epsilon$, the characteristic function of the random variables $\epsilon_{j,m}$, is assumed to be real valued and strictly positive. Further we assume that there exist constants $c_1, c_2 > 0$ with*

$$\phi_\epsilon(\tau) > c_1 e^{-c_2|\tau|}$$

*for $\tau \in I\!R$.*

**A6** *The volatilities $(\xi_t)_{t\in I\!N}$ possess a strictly positive and continuously differentiable stationary density $f_\xi$ with*

$$\left\|f_\xi^{(k)}\right\|_\infty < \infty, \qquad k = 0, 1.$$

**A7** *$U$ denotes a real-valued random variable with density $K$ and real valued characteristic function $\phi_U \in C_C(I\!R)$ with $\operatorname{supp}(\phi_U) \subset [-1, 1]$.*

Now we are going to replace the numerator and the denominator of (2.6) by quantities which do not depend on the unknown characteristic function $\phi_\epsilon$. Let us start with the denominator (2.7), which involves the deconvolution kernel $K_T$, cf. (2.8). Instead of (2.7) we refer to the alternative representation (2.9) in the

following.
In a first step we replace the integral in (2.9) by a Riemann-sum on $[-\lambda_T^{-1}, \lambda_T^{-1}]$ with refinement $d_T$. Secondly, as already mentioned above, we replace $\phi_\epsilon$ by

$$\hat{\phi}_\epsilon^T = \sqrt{\tilde{\phi}_Z^T}, \tag{2.13}$$

where

$$\tilde{\phi}_Z^T(\tau) := \max\left(\mathcal{R}e\left(\hat{\phi}_Z^T(\tau)\right), \frac{c_T^2}{2}\right) \tag{2.14}$$

with

$$\hat{\phi}_Z^T(\tau) := (T(M-1))^{-1}\sum_{j=1}^{T}\sum_{m=2}^{M}\exp(i\tau Z_{j,m}) \tag{2.15}$$

and $c_T$ denoting a lower bound of $\phi_\epsilon$ on $[-\lambda_T^{-1}, \lambda_T^{-1}]$ which will be calculated and defined in (2.20). These steps lead to the estimator

$$\hat{f}_\xi^T(x) := \frac{1}{2\pi}\sum_{j=1}^{S_T} e^{-i\tau_j^T x}\phi_U(\lambda_T\tau_j^T)\frac{\hat{\phi}_X^T(\tau_j^T)}{\hat{\phi}_\epsilon^T(\tau_j^T)}\, d_T \tag{2.16}$$

with

$$\begin{aligned}
0 \; &< \; c < \frac{1}{11c_2}, \\
\lambda_T \; &:= c\,(\log T)^{-1}, \\
a_T \; &:= T^{-3c_2\cdot c}, \\
d_T \; &:= a_T\lambda_T, \\
S_T \; &:= 2(\lambda_T d_T)^{-1} \text{ and} \\
\tau_j^T \; &:= -\frac{1}{\lambda_T} + (j-1)d_T \text{ for } j \in \{1,\ldots,S_T+1\}.
\end{aligned} \tag{2.17}$$

(2.16) can also be written in the following form similar to (2.7)

$$\hat{f}_\xi^T(x) = \frac{1}{TM\lambda_T}\sum_{j=1}^{T}\sum_{m=1}^{M}\hat{K}_T\left(\frac{x - X_{j,m}}{\lambda_T}\right) \tag{2.18}$$

with

$$\hat{K}_T(x) = \frac{1}{2\pi}\sum_{j=1}^{S_T} e^{-it_j^T x}\frac{\phi_U(t_j^T)}{\hat{\phi}_\epsilon^T(t_j^T/\lambda_T)}\,\delta_T, \tag{2.19}$$

where

$$t_j^T := \lambda_T\tau_j^T \quad \text{and} \quad \delta_T := \lambda_T d_T.$$

Assumption A5 and the choice of $\lambda_T$ ensure, that

$$\phi_\epsilon(\tau) > \frac{T^{-c_2 c}}{c_1} =: c_T, \tag{2.20}$$

for $\tau \in [-\lambda_T^{-1}, \lambda_T^{-1}]$.

For the estimator of $f_\xi$, defined in (2.16) and (2.18) we obtain the following result:

**2.2 Theorem.**
*If A1, A2, A5, A6 and A7 are fulfilled and if all tuning parameters are chosen as in* (2.17) *then*

$$\hat{f}_\xi^T(x) \xrightarrow[T\to\infty]{p} f_\xi(x) \quad \forall\, x \in I\!R.$$

The proof is deferred to Appendix A.

Now we have to show, that we can replace the numerator of (2.6) by a quantity which converges to $m(x)f_\xi(x)$ under certain assumptions.
This task is a bit more technical. In order to limit the difficulties we use a sample splitting, for a fixed $a \in (0,1)$ we introduce

$$T_1 = \lfloor aT \rfloor \quad \text{and} \quad T_2 = T - T_1\ , \tag{2.21}$$

where $\lfloor aT \rfloor$ denotes the largest integer not greater than $aT$.
We now estimate the numerator by

$$\frac{1}{T_2 M \lambda_T} \sum_{j=T_1+1}^{T} \sum_{m=1}^{M} \hat{K}_T^1\left(\frac{x - X_{j,m}}{\lambda_T}\right) X_{j+1,m} \tag{2.22}$$

with

$$\hat{K}_T^1(x) := \frac{1}{2\pi} \sum_{j=1}^{S_T} e^{-it_j^T x} \frac{\phi_U(t_j^T)}{\hat{\phi}_\epsilon^{T,1}(t_j^T/\lambda_T)}\, \delta_T, \tag{2.23}$$

where $\hat{\phi}_\epsilon^{T,1}$ is defined just like $\hat{\phi}_\epsilon^T$ but uses only the first part of the observations with length $T_1$.
Thus we get the following estimator of $m$:

$$\hat{m}_T(x) = \frac{\frac{1}{TM\lambda_T} \sum_{j=T_1+1}^{T} \sum_{m=1}^{M} \hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right) X_{j+1,m}}{\hat{f}_\xi^T(x)}. \tag{2.24}$$

For the numerator of this quantity we have

**2.3 Proposition.**
*If A1, A2, A3, A4, A5, A6 and A7 are fulfilled and if all tuning parameters are chosen as in* (2.17) *and in* (2.21), *then*

$$\frac{1}{T_2 M \lambda_T} \sum_{j=T_1+1}^{T} \sum_{m=1}^{M} \hat{K}_T^1 \left( \frac{x - X_{j,m}}{\lambda_T} \right) X_{j+1,m} \xrightarrow[T\to\infty]{p} m(x) f_\xi(x).$$

This proposition - whose proof is also deferred to Appendix A - together with Theorem 2.2 immediately leads to the following consistency result of our estimator.

**2.4 Theorem.**
*If A1, A2, A3, A4, A5, A6 and A7 are fulfilled and if all tuning parameters are chosen as in* (2.17) *and in* (2.21), *then*

$$\hat{m}_T(x) \xrightarrow[T\to\infty]{p} m(x).$$

# Chapter 3

# Estimation in an Errors-in-Variables Model

In chapter 4 two models will be proposed in which the volatility can be observed up to an error, which converges to zero in probability with increasing sample size. Again we will be interested in estimators for the stationary density of the volatility and the autoregression function and their asymptotic behaviour. Therefore we will have to establish a technical framework first, which will be done in the actual chapter.

We assume again, that volatility follows the nonparametric structure (1.6) and furthermore, that for $T \in I\!N$ we can observe $\xi_i$ disturbed by a random noise $\epsilon_i^T$. Let us denote this observable quantity $\hat{\xi}_i^T$, which fulfills

$$\hat{\xi}_i^T = \xi_i + \epsilon_i^T.$$

If we could observe the random-variables $\xi_i$, which follow the autoregressive structure (1.6), we would - as already mentioned in the introduction - estimate $m$ directly by the usual Nadaraya-Watson estimator:

$$\tilde{m}_T(x) = \frac{1}{T\lambda_T} \sum_{i=1}^{T} \frac{K\left(\frac{x-\xi_i}{\lambda_T}\right)\xi_{i+1}}{\tilde{f}_\xi^T(x)} \tag{3.1}$$

with

$$\tilde{f}_\xi^T(x) = \frac{1}{T\lambda_T} \sum_{t=1}^{T} K\left(\frac{x-\xi_i}{\lambda_T}\right), \tag{3.2}$$

where $K$ denotes a probability-density, usually compactly supported and symmetric and $\lambda_T$ a smoothing parameter - the so-called bandwidth.

Since we are not able to observe $\xi_i$ but $\hat{\xi}_i^T$ we replace $\xi_i$ in (3.1) and (3.2), which leads us to the estimators

$$\hat{m}_T(x) := \frac{1}{T\lambda_T} \sum_{i=1}^{T} \frac{K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\hat{\xi}_{i+1}^T}{\hat{f}_\xi^T(x)} \tag{3.3}$$

with

$$\hat{f}_\xi^T(x) = \frac{1}{T\lambda_T} \sum_{i=1}^{T} K\left(\frac{x - \hat{\xi}_i^T}{\lambda_T}\right). \tag{3.4}$$

To achieve our results the following additional assumptions are necessary.

**3.1 Assumptions.**

**A8** *$\epsilon_i^T$, $i \in I\!N$, $T \in I\!N$ are independent random variables and fulfill*

$$E(\epsilon_i^T) = 0 \quad \textit{and} \quad E((\epsilon_i^T)^k) = O\left(\frac{1}{M_T}\right) \quad \textit{for} \quad k = 2, 4, 6.$$

*At this point $M_T$ should just be regarded as a series of natural numbers, fulfilling conditions denoted in A13. In the next chapter $M_T$ will be substantiated to describe the number of observations taken at one day.*

**A9** *$m$ is three times continuously differentiable with*

$$\left\|m^{(k)}\right\|_\infty < \infty, \qquad k = 1, 2, 3.$$

*Notice that again $m$ itself is not necessarily bounded.*

**A10** *The volatilities possess a strictly positive and twice continuously differentiable density $f_\xi$ with*

$$\left\|f_\xi^{(k)}\right\|_\infty < \infty, \qquad k = 0, 1, 2.$$

**A11** *$U$ denotes a symmetric real-valued random variable with density $K \in C_c(I\!R)$ and $\operatorname{supp}(K) = [-a, a]$, $a \in I\!R$, which directly implies the existence of all moments and the existence of $\|K\|_{L_2}$, $\|K\|_\infty$ and $\|K'\|_\infty$.*

**A12** *$\lambda_T$ is chosen such that $T\lambda_T^5 \stackrel{T\to\infty}{\longrightarrow} c_\lambda^2$.*

**A13** *a) $M_T$ is chosen such that $(T^2/M_T^5) \stackrel{T\to\infty}{\longrightarrow} 0$,*

*b) $M_T \stackrel{T\to\infty}{\longrightarrow} \infty$.*

If we choose $M_T$ such that $(T^2/M_T^5) \stackrel{T\to\infty}{\longrightarrow} 0$ (Assumption A13a) and $\lambda_T$ such that $(\lambda_T/T^5) \stackrel{T\to\infty}{\longrightarrow} 0$, we can achieve the same convergence results as in the situation where $\xi_i$ is observable (cf. Franke, Kreiss, Mammen (2002) [49] for the asymptotics in case that the daily mean volatilities $\xi_i$ are directly observable). This is the main result of this chapter and is formulated in the following Theorem.

**3.2 Theorem.**
*If A1, A2, A3, A8, A9, A10, A11, A12 and A13a are fulfilled, then*

$$\begin{aligned}
&\sqrt{T\lambda_T}\left(\hat{m}_T(x) - m(x)\right) \\
&\quad = \frac{\frac{1}{\sqrt{T\lambda_T}} \sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\left(\hat{\xi}_{i+1}^T - m(x)\right)}{\frac{1}{T\lambda_T}\sum_{t=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)} \\
&\quad \xrightarrow[T\to\infty]{d} \mathcal{N}\left(c_\lambda \frac{\mathrm{E}\left(U^2\right)\left(-m'(x)f_\xi'(x) + \frac{1}{2} f_\xi(x) m''(x)\right)}{f_\xi(x)}, \frac{\sigma_\eta^2 \left\|K\right\|_{L_2}^2}{f_\xi(x)}\right).
\end{aligned} \tag{3.5}$$

Aside from this we still achieve consistency of the estimator $\hat{m}_T$ if we only have $M_T \overset{T\to\infty}{\longrightarrow} \infty$ (Assumption A13b) instead of $(T^2/M_T^5) \overset{T\to\infty}{\longrightarrow} 0$ (Assumption A13a).

**3.3 Theorem.**
*If A1, A2, A3, A8, A9, A10, A11, A12 and A13b are fulfilled, then*

$$\hat{m}_T(x) \xrightarrow[T\to\infty]{p} m(x).$$

The proofs of Theorem 3.2 and Theorem 3.3 are deferred to Appendix B.2.

Another result - which will be necessary for the proof of the Theorems above - is the consistency of the estimator of $f_\xi$ defined in (3.4) and reads as follows:

**3.4 Theorem.**
*If A1, A2, A8, A10, A11, A12 and A13b are fulfilled, then*

$$\hat{f}_\xi^T(x) = \frac{1}{T\lambda_T}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right) \xrightarrow[T\to\infty]{p} f_\xi(x).$$

The proof of this Theorem can be found in Appendix B.1.

# Chapter 4

# Estimators based on higher frequent Data

In this chapter we consider some financial time series $(S_t)$, which is observable reasonably often during a daily observation period so that with increasing time-horizon $T$ we can take an increasing number of observations at each day. More exactly this means that for $T \in I\!N$, $i \in \{0, \ldots, T+1\}$ and $j \in \{0, \ldots, M_T\}$ (where $M_T \stackrel{T\to\infty}{\longrightarrow} \infty$) we denote by $S^T_{i,j}$, the j-th observation of the i-th day. The $(M_T + 1)$ observations of each day are assumed to be homogenously distributed over the daily observation period. Notice that the superscript T just indicates how many days $(T + 2)$ and how often at each day $(M_T + 1)$ we observe the time series.
In practical use one should take care not to take too many observations, to avoid the data to be contaminated by short term trading phenomena, such as the effects of the bid-ask spread. This problem of so-called market microstructure noise has already been mentioned and discussed in Cox and Rubinstein (1989) [30] and Brown (1990) [24]. Sources of market microstructure noise are presented in Black (1976) [14], Amihud and Mendelsohn (1987) [2] and Harris (1990, 1991) [54, 55]. Presently volatility estimation in the presence of market microstructure noise is a very active area of research, which was initiated by Zhou (1996) [107]. In Aït Sahalia, Mykland and Zhang (2005) [1] the authors discuss the question of "How Often to Sample a Continous-Time Process in the Presence of Market Microstructure Noise". Their concrete implication for empirical work with high frequency data is to sample less frequently. A typical choice was 5 minutes and up. Aït Sahalia et al. mainly deal with models with constant volatility. In Zhang, Mykland and Aït Sahalia (2005) [106] the authors propose - in their framework - a way to find the optimal sampling frequency in a model with stochastic volatility. Recent work on market microstructure noise can also be found in Hansen and Lunde (2005) [53] and in Hasbrouck's lecture-notes to appear in 2007 [56].

## 4.1 An Integrated Volatilities Autoregression Model

For $T \in I\!N, i \in \{1, \ldots, T+1\}, j \in \{1, \ldots, M_T\}$ we define the intra-day log-returns by

$$L_{i,j}^T = \log\left(S_{i,j}^T\right) - \log\left(S_{i,j-1}^T\right). \tag{4.1}$$

This means, that $L_{i,j}^T$ describes (an approximation of) the relative change between direct neighbour observations and for each $i$ summing up all $L_{i,j}$, $j = 1, \ldots, M_T$ yields approximately the day's return. We assume, that

$$L_{i,j}^T = \frac{\mu}{M_T} + \sigma_i \kappa_{i,j}^T \eta_{i,j}^T, \tag{4.2}$$

with $\sigma_i = \exp(\xi_i)$, $\xi_i$ following the autoregressive structure

$$\xi_i = m(\xi_{i-1}) + \eta_i, \qquad i = 1, 2, \ldots\ , \tag{4.3}$$

and $\kappa_{i,j}^T$ and $\eta_{i,j}^T$ fulfilling the following assumptions:

**4.1 Assumptions.**

**A14** *The random variables $\eta_{i,j}^T$, $T \in I\!N, i \in \{1, \ldots, T+1\}, j \in \{1, \ldots, M_T\}$, are i.i.d., centered, of variance 1, independent of $\sigma_i$ and their moments up to order 12 exist.*

**A15** *$\kappa_{i,j}^T$, $T \in I\!N, i \in \{1, \ldots, T+1\}, j \in \{1, \ldots, M_T\}$, are deterministic with*

$$\sum_{j=1}^{M_T} \left(\kappa_{i,j}^T\right)^2 = 1$$

*and*

$$\mathrm{E}\left(\left[\sum_{j=1}^{M_T} \left(\kappa_{i,j}^T\right)^2 \left(\left(\eta_{i,j}^T\right)^2 - 1\right)\right]^k\right) = O\left(\frac{1}{M_T}\right) \text{ for } k = 2, 4, 6. \tag{4.4}$$

Notice, that under these assumptions all random effects in (4.1) are covered by the $\sigma$s and $\eta^T$s and that A15 and especially (4.4) are e.g. fulfilled if A14 is fulfilled and the $\kappa_{i,j}^T$ are chosen as ${M_T}^{-0.5}$.

These assumptions ensure, that

$$\mathrm{E}\left(\sum_{i=1}^{M_T} \kappa_{i,j}^2 \eta_{i,j}^2\right) = \sum_{i=1}^{M_T} \kappa_{i,j}^2 \mathrm{E}\left(\eta_{i,j}^2\right) = 1 \tag{4.5}$$

and

$$\begin{aligned}\operatorname{Var}\left(\sum_{i=1}^{M_T}\kappa_{i,j}^2\eta_{i,j}^2\right) &= \operatorname{Var}\left(\sum_{i=1}^{M_T}\kappa_{i,j}^2\eta_{i,j}^2-1\right) = \operatorname{Var}\left(\sum_{i=1}^{M_T}\kappa_{i,j}^2\eta_{i,j}^2-\sum_{i=1}^{M_T}\kappa_{i,j}^2\right)\\ &= \operatorname{Var}\left(\sum_{i=1}^{M_T}\kappa_{i,j}^2\left[\eta_{i,j}^2-1\right]\right) \leq \mathrm{E}\left(\left(\sum_{i=1}^{M_T}\kappa_{i,j}^2\left[\eta_{i,j}^2-1\right]\right)^2\right)\\ &= O\left(\frac{1}{M_T}\right),\end{aligned}$$

which implies

$$\sum_{i=1}^{M_T}\left(\kappa_{i,j}^T\right)^2\left(\eta_{i,j}^T\right)^2 \xrightarrow[T\to\infty]{p} 1, \tag{4.6}$$

and

$$\sum_{i=1}^{M_T}\left(L_{i,j}^T-\frac{\mu}{M_T}\right)^2 \xrightarrow[T\to\infty]{p} \sigma_i^2. \tag{4.7}$$

Notice, that under the assumptions A14 and A15 also

$$\sum_{i=1}^{M_T}\left(L_{i,j}^T\right)^2 \xrightarrow[T\to\infty]{p} \sigma_i^2, \tag{4.8}$$

which can be shown quite easily.
We are thus in the same situation as in continous stochastic volatility-models, which have been presented in the introduction in equation (1.4). In these models quadratic variation of the log-price process of an asset is equal to the so-called integrated volatility

$$\sigma^{2\star}(t) = \int_0^t \sigma^2(u)\,du,$$

which has been pointed out e.g. in Andersen and Bollerslev (1998) [3], Comte and Renault (1998) [27] and Barndorff-Nielsen and Shephard (2001, 2002) [10, 11]. This equality is independent of the model of the instantaneous volatility $\sigma(t)$ and the drift-term in (1.4). Since quadratic variation of a process $x(s)$ between 0 and $t$, denoted by $[x](t)$, is defined as

$$[x](t) := p - \lim_{q\to\infty}\sum_{i=1}^{q}\left(x(t_{i+1}^q)-x(t_i^q)\right)^2,$$

for any sequence of of partitions $0 = t_0^q < t_1^q < \ldots < t_q^q = t$ with $\sup_{i=1,\ldots,q-1}\left(t_{i+1}^q - t_i^q\right) \xrightarrow{q\to\infty} 0$, the sum of squared log-returns of a price-process

fulfilling the SDE (1.4), usually called realised volatility, converges towards the integrated volatility, when the number of observations tends to infinity and the distance between the observations tends to zero. Recent work on so-called quadratic variation estimation has e.g. been done by Andersen, Bollerslev, Diebold and Labys (2001) [5] and Maheu and McCurdy (2001) [67] in foreign exchange markets and Andersen, Bollerslev, Diebold and Ebens (2001) [4] and Areal and Taylor (2002) [7] in equity markets. A theoretical comparison between integrated and realised volatility and some results - concerning the presence of leverage effects and time-varying drift - complementing those of Barndorff-Nielsen and Shephard (2002) can be found in Meddahi (2002) [73]. A generalisation of quadratic variation estimation - so-called power variation - in stochastic volatility models has been investigated in Barndorf-Nielsen, Graversen and Shephard (2003) [9] and Barndorf-Nielsen, Shephard (2003) [12].
Since the sum of intra-day log-returns of a process described by our model given in (4.2) converges towards $\sigma_i^2$, as already highlighted above, we can identify our random variable $\sigma_i$ with the square-root of the i-th day integrated volatility and will therefore name the random variables $\sigma_i^2$ integrated volatility as well.
Thus in this model-setting the autoregression (4.3) describes a dependence structure between integrated volatilities and it is the aim of this section to provide an estimator of the autoregression function $m$ of a specific logarithmic transformation of the integrated volatility.
Defining $X_{i,j}^T = \left(L_{i,j}^T - \frac{\mu}{M_T}\right)^2$, we get:

$$\begin{aligned}
\frac{1}{2}\log\left(\sum_{i=1}^{M_T} X_{i,j}^T\right) &= \frac{1}{2}\log\left(\sum_{i=1}^{M_T}\sigma_i^2\left(\kappa_{i,j}^T\right)^2\left(\eta_{i,j}^T\right)^2\right) \\
&= \frac{1}{2}\log\left(\sigma_i^2\right) + \frac{1}{2}\log\left(\sum_{i=1}^{M_T}\left(\kappa_{i,j}^T\right)^2\left(\eta_{i,j}^T\right)^2\right) \\
&= \xi_i + \frac{1}{2}\left(\sum_{i=1}^{M_T}\left(\kappa_{i,j}^T\right)^2\left(\eta_{i,j}^T\right)^2 - 1\right) \\
&\quad + \frac{1}{4}\left(\sum_{i=1}^{M_T}\left(\kappa_{i,j}^T\right)^2\left(\eta_{i,j}^T\right)^2 - 1\right)^2 \frac{1}{\theta},
\end{aligned} \tag{4.9}$$

where $\theta$ denotes an appropriate value between 1 and $\sum_{i=1}^{M_T}\left(\kappa_{i,j}^T\right)^2\left(\eta_{i,j}\right)^2$. Because of (4.6) and a faster convergence rate of the third summand we neglect this term and simplify (4.9) to

$$\frac{1}{2}\log\left(\sum_{i=1}^{M_T} X_{i,j}^T\right) = \xi_i + \frac{1}{2}\left(\sum_{i=1}^{M_T}\left(\kappa_{i,j}^T\right)^2\left[\left(\eta_{i,j}^T\right)^2 - 1\right]\right). \tag{4.10}$$

Denoting the second addend on the right-hand-side as $\epsilon_i^T$, it is quite obvious, that these random variables fulfill condition A8.
The independence of the $\epsilon_i^T$s is ensured by A14, $\mathrm{E}\left(\epsilon_i^T\right) = 0$ follows directly from (4.5) and $\mathrm{E}\left(\left(\epsilon_i^T\right)^k\right) = O\left(\frac{1}{M_T}\right)$ from (4.4). So in this situation we can define the estimator of $\xi_i$ by

$$\hat{\xi}_i^T = \frac{1}{2}\log\left(\sum_{i=1}^{M_T} X_{i,j}\right)$$

and make use of the results given in chapter 3.

## 4.2 Daily Mean Stochastic Volatility Model

For $i, j \geq 1$ we define the inter-day return $R_{i,j}^T$ by

$$R_{i,j}^T = \frac{S_{i,j}^T - S_{i-1,j}^T}{S_{i-1,j}^T}.$$

These returns describe the relative price-change between two observations taken at the same time on two consecutive days,
We assume that

$$R_{i,j}^T = \mu + \sigma_{i,j}^T V_{i,j}^T \tag{4.11}$$

holds with

$$\sigma_{i,j}^T = \exp(\xi_{i,j}^T)$$

and i.i.d. random variables $V_{i,j}^T$. Furthermore we assume that there exists a non-observable so-called daily mean volatility $\xi_i$, which is supposed to follow the nonparametric autoregressive structure

$$\xi_i = m(\xi_{i-1}) + \eta_i, \qquad i = 1, 2, \dots \ , \tag{4.12}$$

where $(\eta_i)_{i \in \mathbb{N}}$ denotes a family of i.i.d. zero-mean random variables.
The volatilities $\xi_{i,j}^T$, which correspond to the returns $R_{i,j}^T, j = 1, \dots, M_T$ , from a single observation period (a day, say), may deviate from the daily mean volatility $\xi_i$ by a random quantity $\zeta_{i,j}^T$ , i.e.

$$\xi_{i,j}^T = \xi_i + \zeta_{i,j}^T \ , \ j = 1, \dots, M_T \ . \tag{4.13}$$

The average deviation

$$\zeta_i^T := \frac{1}{M_T} \sum_{j=1}^{M_T} \zeta_{i,j}^T$$

for a single observation period from the daily mean volatility $\xi_i$ will be assumed to converge to zero in probability with increasing sample size (cf. assumption A8 below). Moreover we need to assume that (again compare A8) the deviations $\zeta_{i,j}^T$ from the daily mean volatility for different observation periods are independent. In order to ensure this in applications it is recommended to separate the observation periods of different days by a sufficiently large time gap.
The conditions claimed to the random variables $\zeta_{i,j}^T$ are given by the following assumption:

**4.2 Assumptions.**

**A16** *$\zeta_{i,j}^T$ is independent of $\zeta_{k,l}^T$, when $i \neq k$ and $\zeta_i^T$ fulfills*

$$E\left(\zeta_i^T\right) = 0 \quad \textit{and} \quad E\left((\zeta_i^T)^k\right) = O\left(\frac{1}{M_T}\right) \quad \textit{for} \quad k = 2, 4, 6.$$

Notice that we do not assume the $\zeta_{i,j}^T$s , $j = 1, \ldots, M_T$, to be independent. But we assume independence between $\zeta_{i,j}^T$ and $\zeta_{k,l}^T$ if $i \neq k$, which ensures independence of the $\zeta_i^T$s, $i = 1, \ldots, T+1$ .

Again we define

$$X_{i,j}^T := \frac{1}{2}\log(R_{i,j}^T - \mu)^2, \qquad \epsilon_{i,j}^T := \frac{1}{2}\log\left(V_{i,j}^T\right)^2$$

and assume $E\left(\epsilon_{i,j}^T\right) = 0$, which can be justified just as in chapter 2.
Since

$$X_{i,j}^T = \xi_{i,j}^T + \epsilon_{i,j}^T$$

we get

$$\begin{aligned} \frac{1}{M_T}\sum_{j=1}^{M_T} X_{i,j}^T &= \frac{1}{M_T}\sum_{j=1}^{M_T}\xi_{i,j}^T + \frac{1}{M_T}\sum_{j=1}^{M_T}\epsilon_{i,j}^T \\ &= \xi_i + \frac{1}{M_T}\sum_{j=1}^{M_T}\zeta_{i,j}^T + \frac{1}{M_T}\sum_{j=1}^{M_T}\epsilon_{i,j}^T. \end{aligned}$$

It is obvious, that the i.i.d. zero mean valued random variables $\epsilon_i^T$ defined by

$$\epsilon_i^T := \frac{1}{M_T}\sum_{j=1}^{M_T}\zeta_{i,j}^T + \frac{1}{M_T}\sum_{j=1}^{M_T}\epsilon_{i,j}^T$$

fulfill condition A8. Thus we can define the observable quantity $\hat{\xi}_i^T$ in this situation by

$$\hat{\xi}_i^T := \frac{1}{M_T}\sum_{j=1}^{M_T} X_{i,j}^T = \xi_i + \epsilon_i^T.$$

Finally we would like to formulate our model in a modified way.
Recall Equation (4.11) and (4.13). Plugging (4.13) into (4.11) we get

$$R_{i,j}^T = \mu + \exp(\xi_i)\exp(\zeta_{i,j}^T)V_{i,j}^T.$$

Denoting

$$\sigma_i := \exp(\xi_i) \quad \text{and} \quad Y_{i,j}^T := \exp(\zeta_{i,j}^T)V_{i,j}^T$$

we get the model

$$R_{i,j}^T = \mu + \sigma_i Y_{i,j}^T \tag{4.14}$$

with a constant volatility over the day $i$. But now the innovation random variables $Y_{i,j}^T$ are not necessarily independent anymore. If we change our assumption A16 into

**A16'** $\log(Y_{i,j}^T)$ is independent from $\log(Y_{k,l}^T)$ for $i \neq k$ and if $Y_i^T := \exp(-M_T) \prod_{j=1}^{M_T} Y_{i,j}^T$ fulfills

$$\mathrm{E}\left(\log\left(Y_i^T\right)\right) = 0 \quad \text{and} \quad \mathrm{E}\left(\log\left(Y_i^T\right)^k\right) = O\left(\frac{1}{M_T}\right) \quad \text{for} \quad k = 2, 4, 6,$$

and define $\hat{\xi}_i^T$ again by

$$\hat{\xi}_i^T := \frac{1}{M_T} \sum_{j=1}^{M_T} X_{i,j}^T$$

the noise term fulfills condition A8 as well.

# Chapter 5

# Simulations

The following estimations result from simulated data, generated by the following mechanism

$$\xi_{i+1} = m(\xi_i) + \eta_i$$

with

$$m(x) = \left[\sqrt{3(x+7)} - 7\right] \cdot 1_{(-\infty,-7)} + \left[\frac{x}{2} - \frac{7}{2}\right] \cdot 1_{[-7,\infty)},$$

where $\xi_0 = 0$ and $\eta_i \sim \mathcal{N}(0, 0.81), i = 1, \ldots, 67000$. For each $i$ an estimator of $\xi_i$ was simulated by adding two noise terms $\zeta_i$ and $\epsilon_i$ to $\xi_i$ and these disturbed random variables were used to estimate the stationary density of $\xi_i$ and the autoregression function $m$, as described in chapter 3. In figures 5.1 - 5.6 the estimated density of $\xi_i$ can always be found on the top and below a plot of the real autoregression function $m$ and the estimated autoregression function.
In the context of our daily-mean-volatility model $\epsilon_i$ represents the deviation caused by the sum of logarithms of the driving process and $\zeta_i$ represents the deviation from the mean of the observed volatilities from $\xi_i$. In the context of the model presented in section 4.1 one should just regard the sum of these two random variables as $\left(\sum_{i=1}^{M_T} \left(\kappa_{i,j}^T\right)^2 \left[\left(\eta_{i,j}^T\right)^2 - 1\right]\right)/2$.
In all estimations $\zeta_i$ is normally distributed with mean zero. In figures 5.1 - 5.3 the standard-deviation is 0.6, in figures 5.4 - 5.6 the standard-deviation is 0.3. In figure 5.1 and figure 5.4 $\epsilon_i$ is also normally distributed with mean zero and standard deviation 0.7 and 0.5 respectively. In the other four cases noise-terms with fatter tails than the normal distribution were used. In figure 5.2 and figure 5.5 each $\epsilon_i$, $i = 1, \ldots, 67000$, is the mean of 10 $t(1)$-distributed random variables scaled to square-mean $0.7^2$ and $0.5^2$, respectively. In figure 5.3 and figure 5.6 for each $\epsilon_i$ the mean of 10 independent Exp(1)-distributed random-variables were subtracted from the mean of 10 other independent random variables. The resulting random variables were scaled to square-mean $0.7^2$ and $0.5^2$, too.
The estimators in figures 5.4 - 5.6 were calculated by using the S-Plus kernel

"Box", the estimators in figures 5.4 - 5.6 by using the S-Plus kernel "Parzen", in all six estimations the bandwidth is chosen as 0.1. In all six cases the estimations of the autoregression function fit quite well to the real function, but we can always see, that on the left-hand-side the regression function tends to be under-estimated, while on the right-hand-side the regression function tends to be over-estimated. In linear regression the so called "attenuation effect", which says, that in the case of an errors-in-variables problem the slope tends to be under-estimated, is well known. Staudenmayer and Ruppert (2004) [90] showed, that a similar result holds true for nonparametric regression as well, which is supported by our estimations.

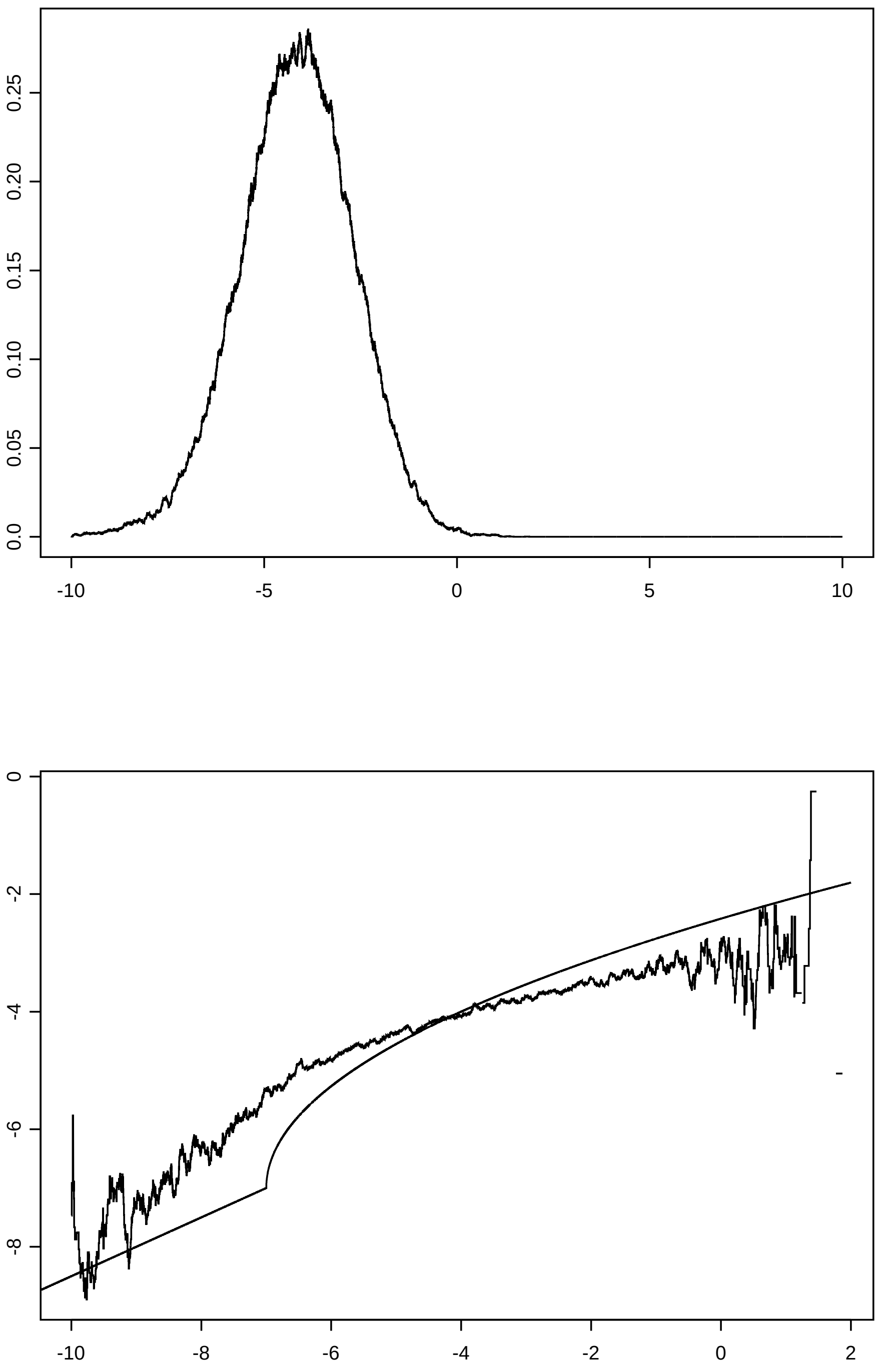



Figure 5.1:

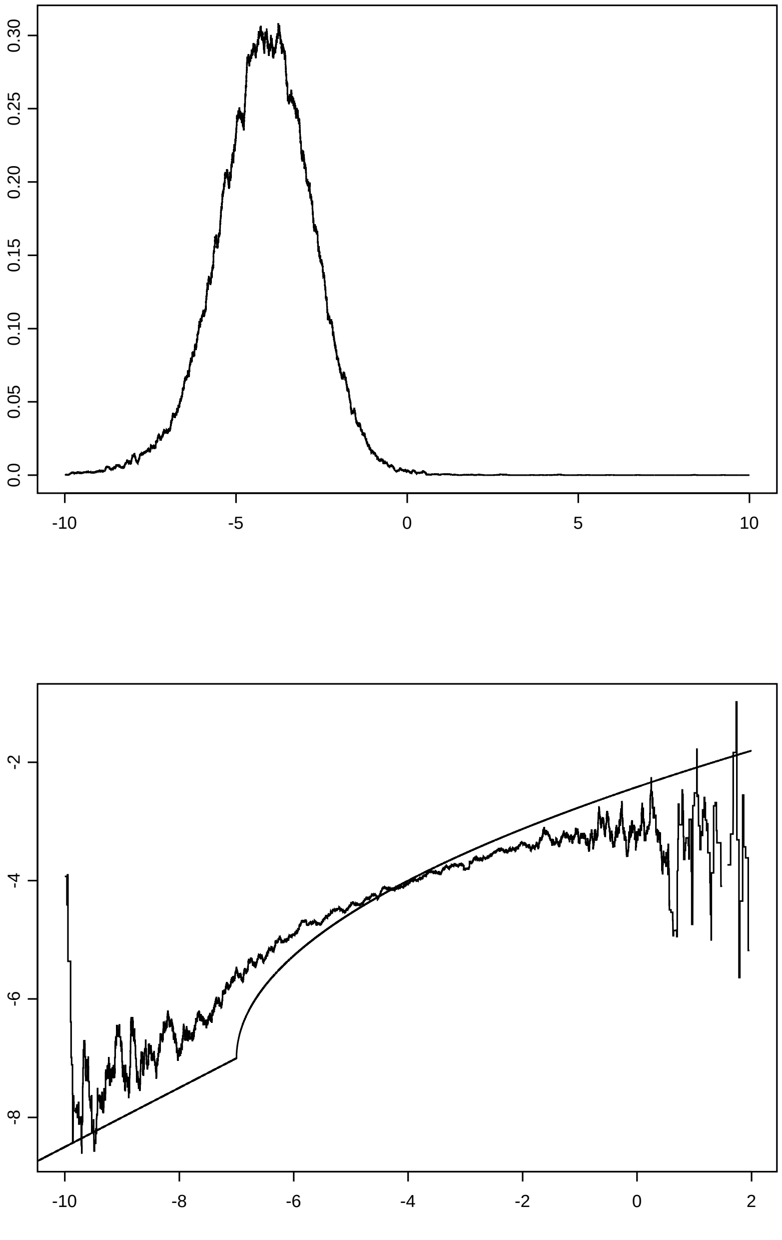
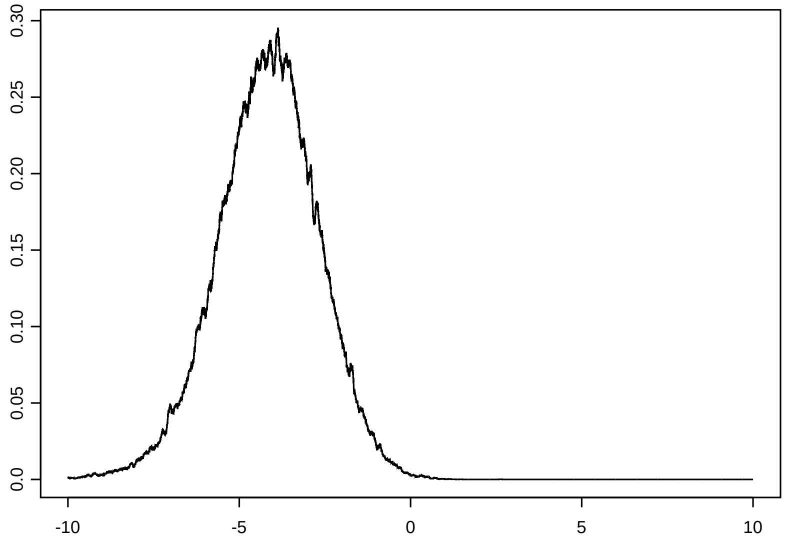



Figure 5.2:



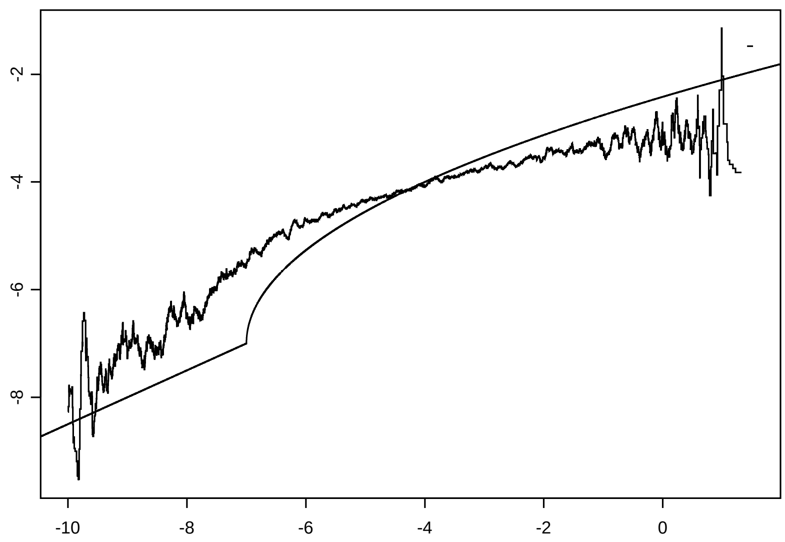




Figure 5.3:

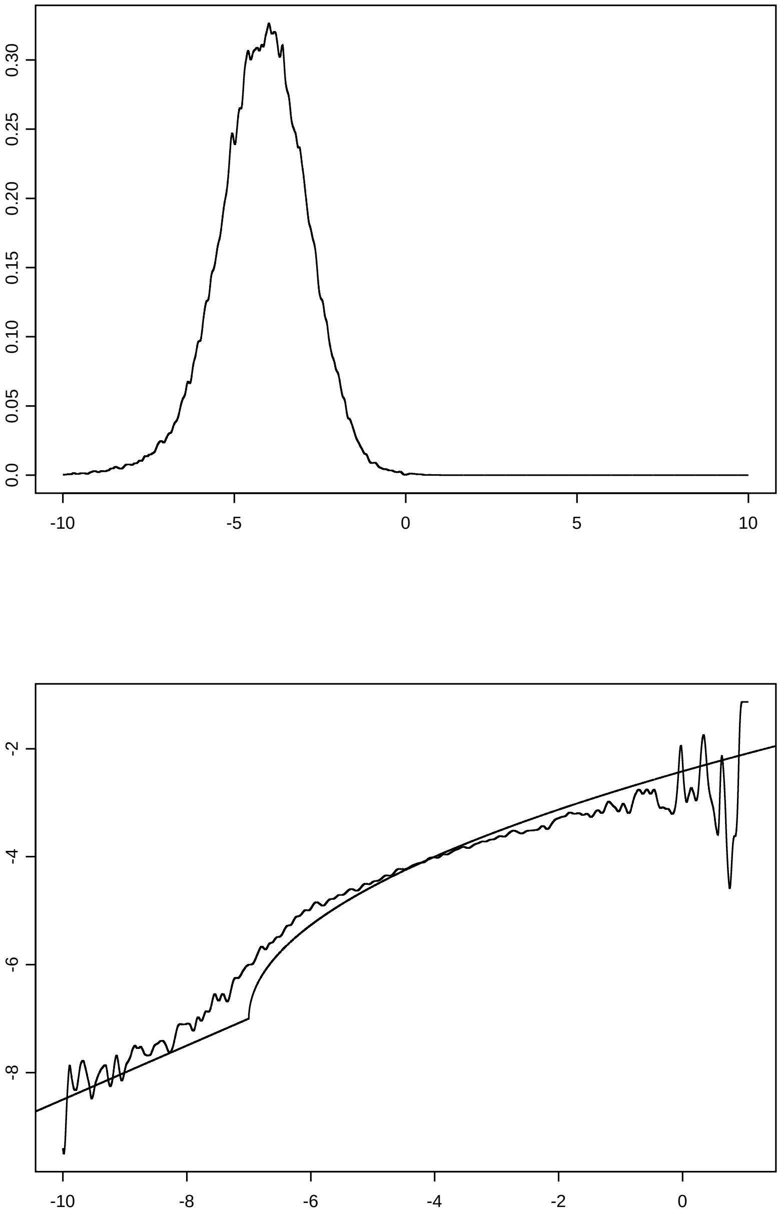



Figure 5.4:

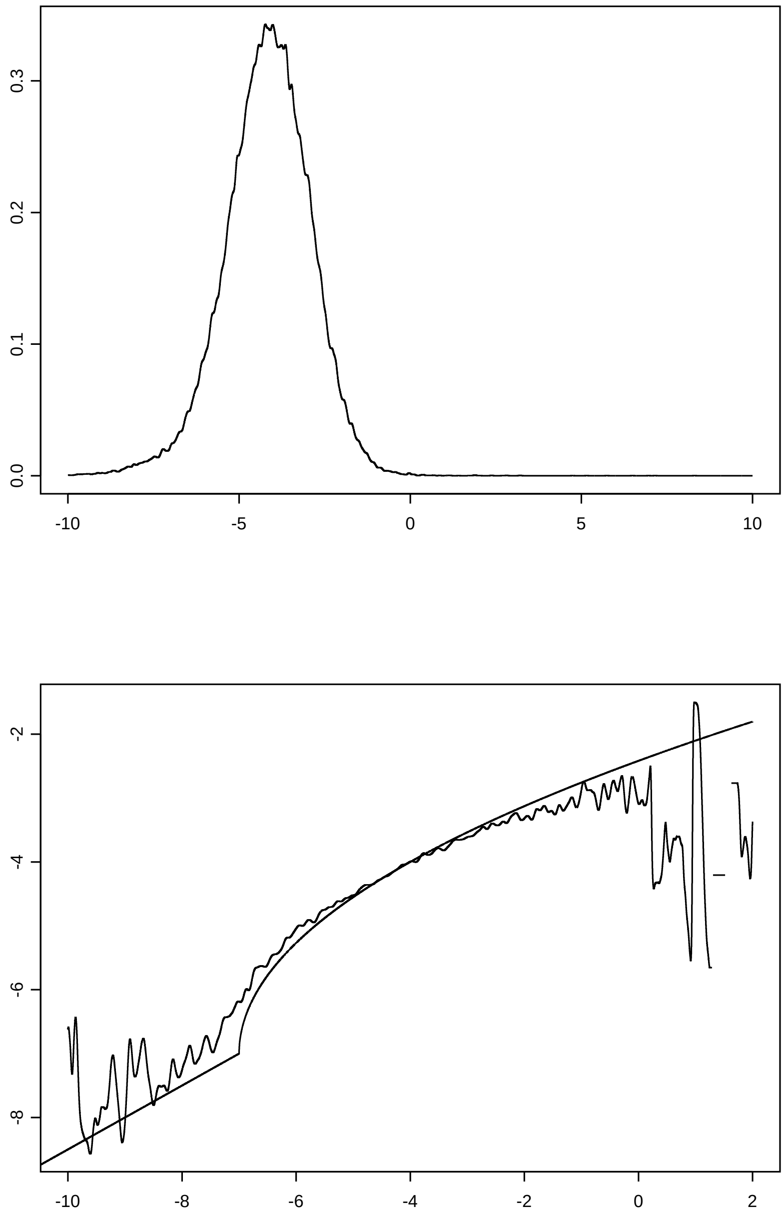
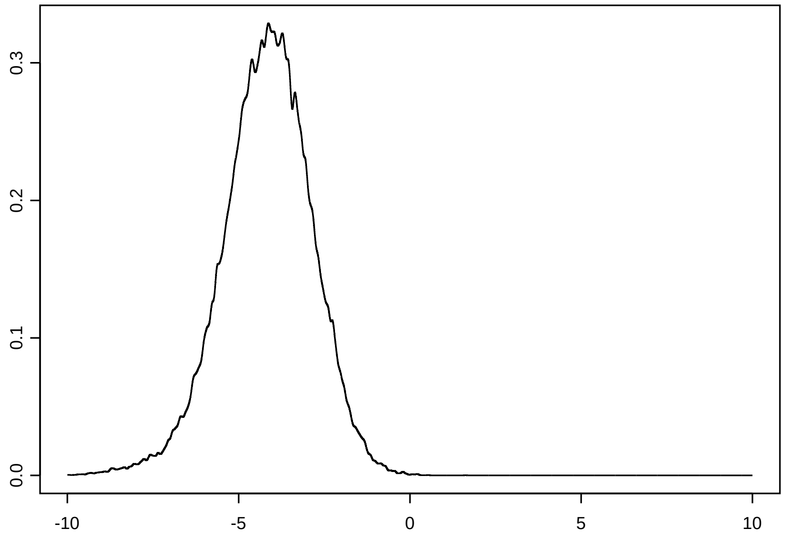



Figure 5.5:



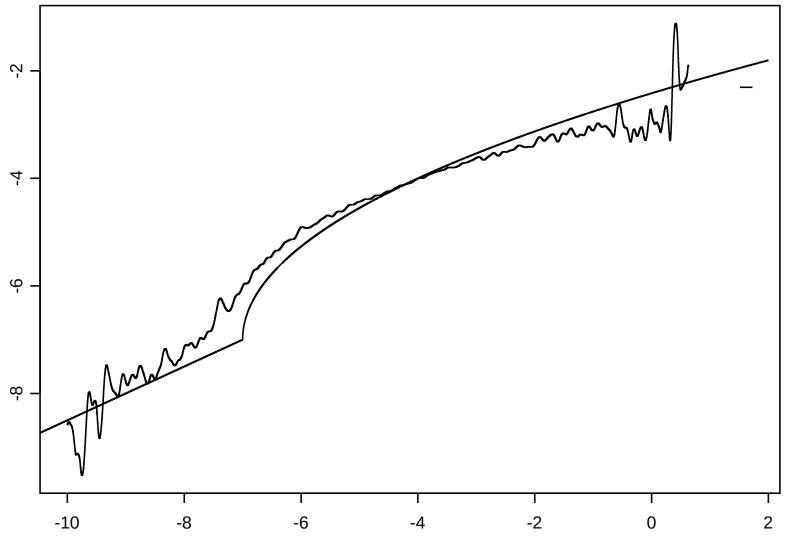




Figure 5.6:

# Chapter 6

# Estimation from Dax Data 1997

In this chapter some estimations of the stationary density of the volatility and the autoregression function of the volatility taken from the 1997 DAX data can be found. The data were observed between 10:40 to 12:00 and used with a distance of 5 and 10 minutes. Figures 6.1 to 6.4 are estimations of the density of the daily-mean volatility and the regression function of the daily mean volatility, i.e. inter-day returns were considered. Figures 6.5 to 6.8 are estimations of the density of the integrated volatility and the regression function of the integrated volatility based on intra-day log-returns, as described in section 4.1. Again the estimator of the stationary density can be be found atop of each page and the estimator of the regression function below. The S-Plus kernels "Box" and "Parzen" and the bandwidths 0.2 and 0.4 were used - the actual settings are always documented within the figures.
In the "daily-mean-volatility"-case, because of the estimations one would assume, that the regression function is almost constant, which would mean, that the volatilities are independent - not a very convincing result. In the "integrated-volatility"-case, there is a certain dependence-structure, which becomes apparent. Between $-6.5$ and $-5.0$ the estimated funtion almost seems to be linear, approximately $m(x) = 0.3 \cdot x - 4$, which is quite more realistic.

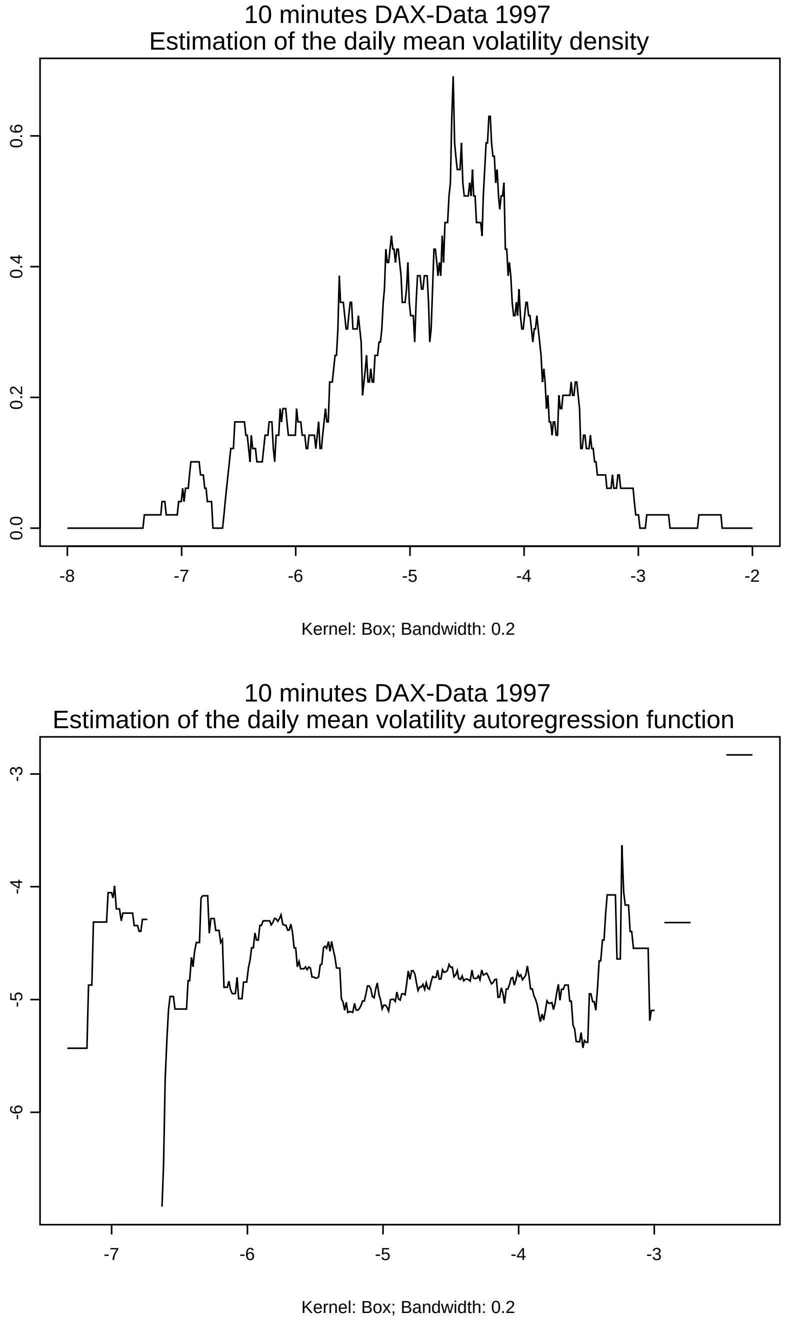

10 minutes DAX-Data 1997
Estimation of the daily mean volatility density
0.0
0.2
0.4
0.6
-8
-7
-6
-5
-4
-3
-2
Kernel: Box; Bandwidth: 0.2
10 minutes DAX-Data 1997
Estimation of the daily mean volatility autoregression function
-3
-4
-5
-6
-7
-6
-5
-4
-3
Kernel: Box; Bandwidth: 0.2

Figure 6.1:

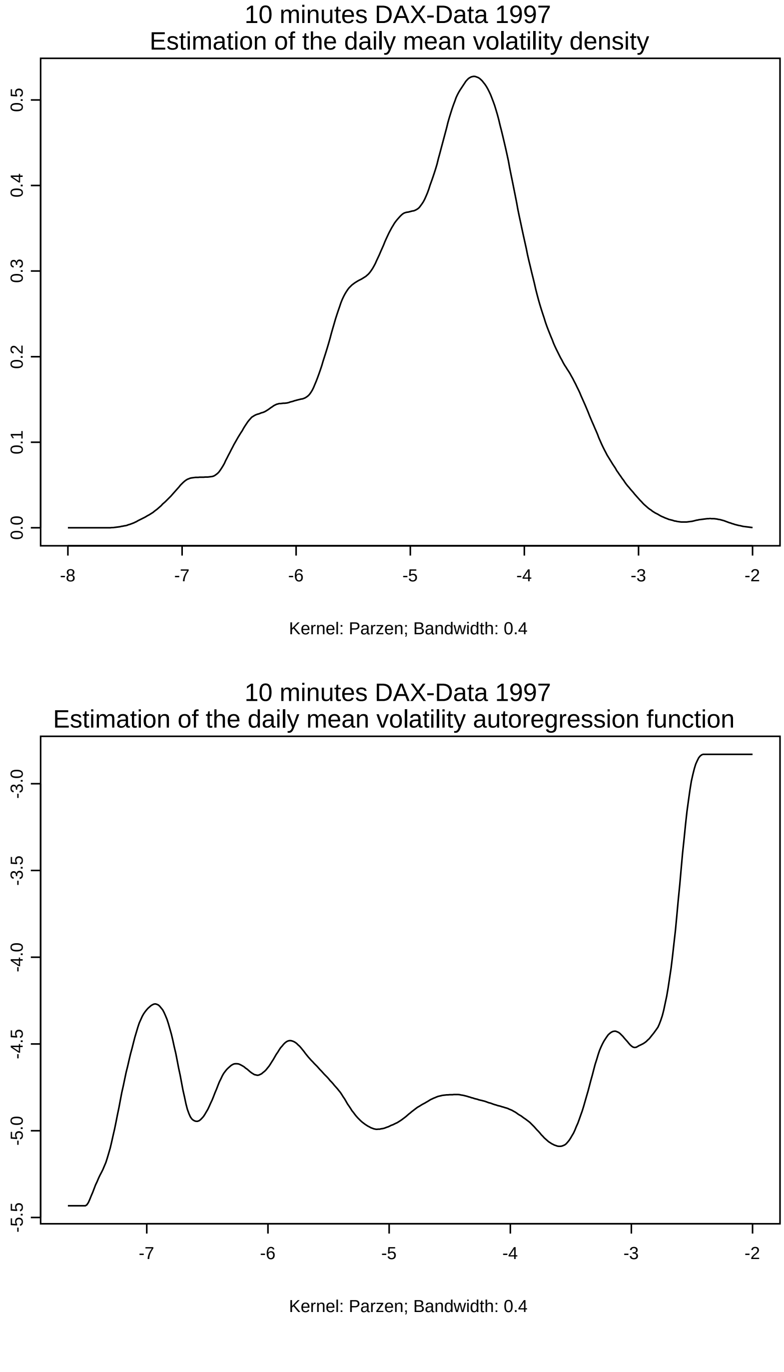
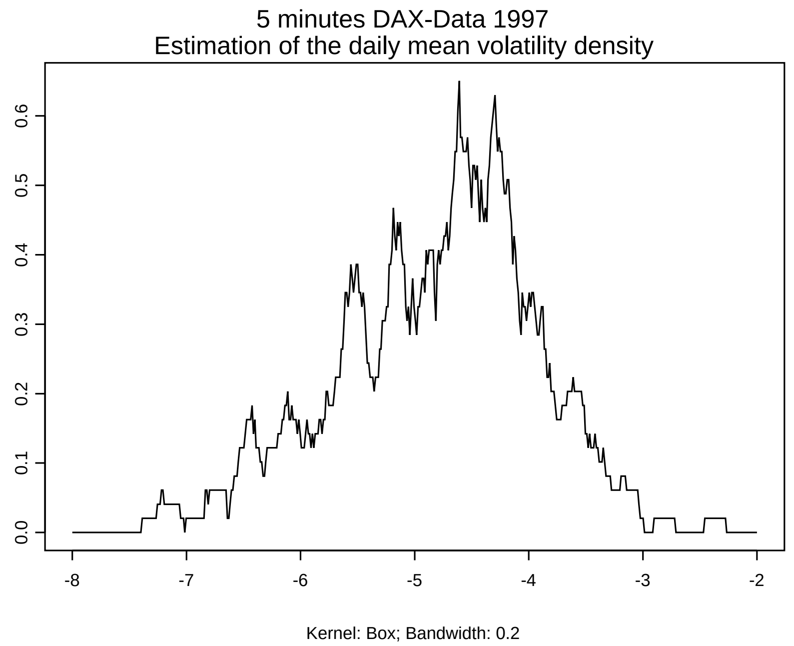

10 minutes DAX-Data 1997
Estimation of the daily mean volatility density
0.0
0.1
0.2
0.3
0.4
0.5
-8
-7
-6
-5
-4
-3
-2
Kernel: Parzen; Bandwidth: 0.4
10 minutes DAX-Data 1997
Estimation of the daily mean volatility autoregression function
-5.5
-5.0
-4.5
-4.0
-3.5
-3.0
-7
-6
-5
-4
-3
-2
Kernel: Parzen; Bandwidth: 0.4

Figure 6.2:


5 minutes DAX-Data 1997
Estimation of the daily mean volatility density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
-8
-7
-6
-5
-4
-3
-2
Kernel: Box; Bandwidth: 0.2

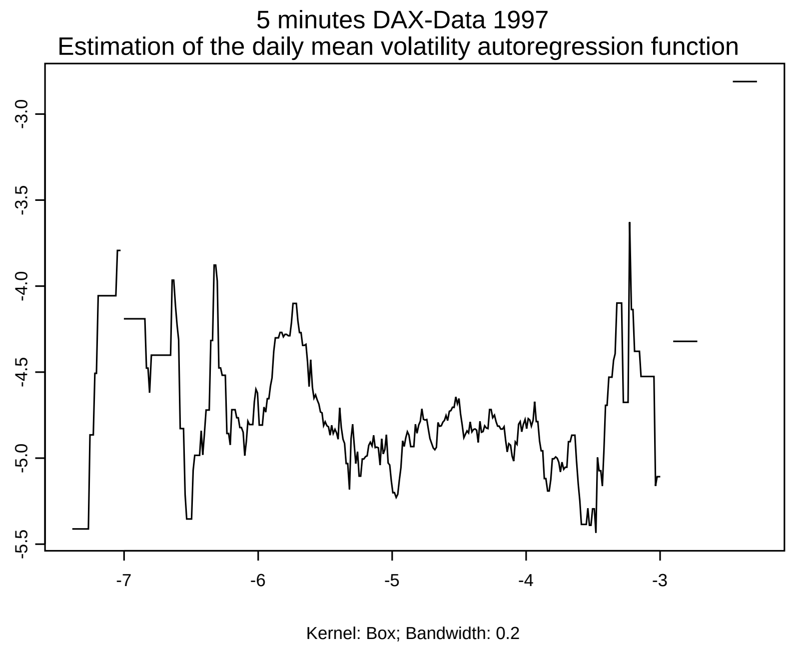
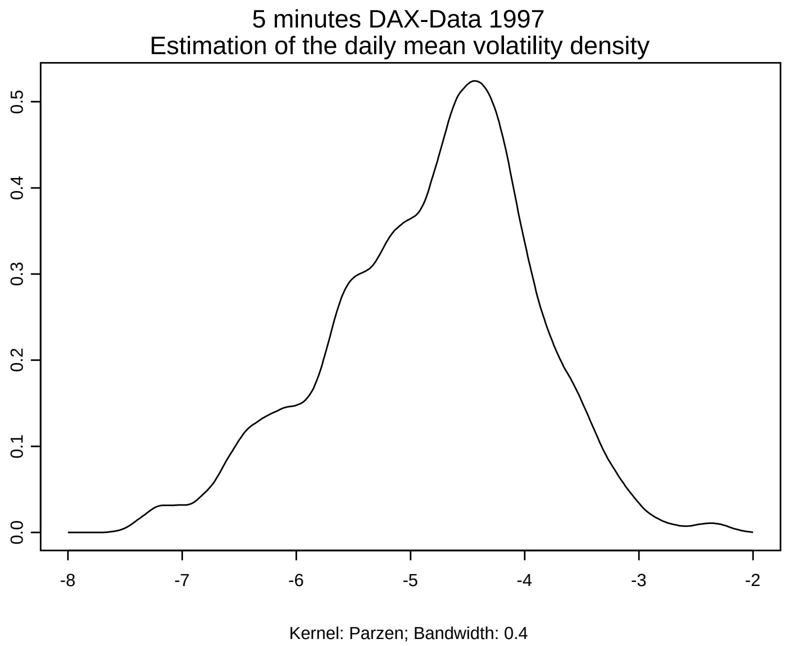

5 minutes DAX-Data 1997
Estimation of the daily mean volatility autoregression function
-3.0
-3.5
-4.0
-4.5
-5.0
-5.5
-7
-6
-5
-4
-3
Kernel: Box; Bandwidth: 0.2

Figure 6.3:


5 minutes DAX-Data 1997
Estimation of the daily mean volatility density
0.0
0.1
0.2
0.3
0.4
0.5
-8
-7
-6
-5
-4
-3
-2
Kernel: Parzen; Bandwidth: 0.4

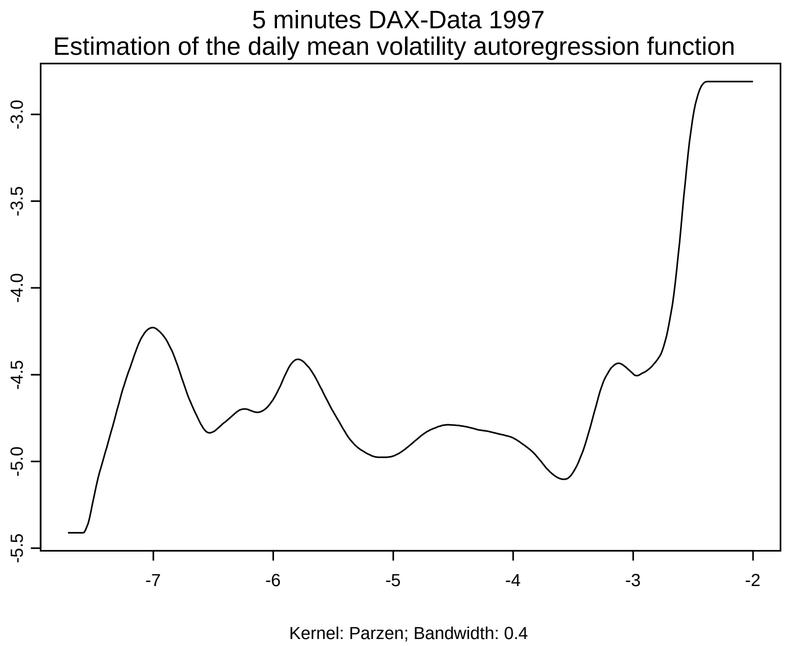

5 minutes DAX-Data 1997
Estimation of the daily mean volatility autoregression function
-5.5
-5.0
-4.5
-4.0
-3.5
-3.0
-7
-6
-5
-4
-3
-2
Kernel: Parzen; Bandwidth: 0.4

Figure 6.4:

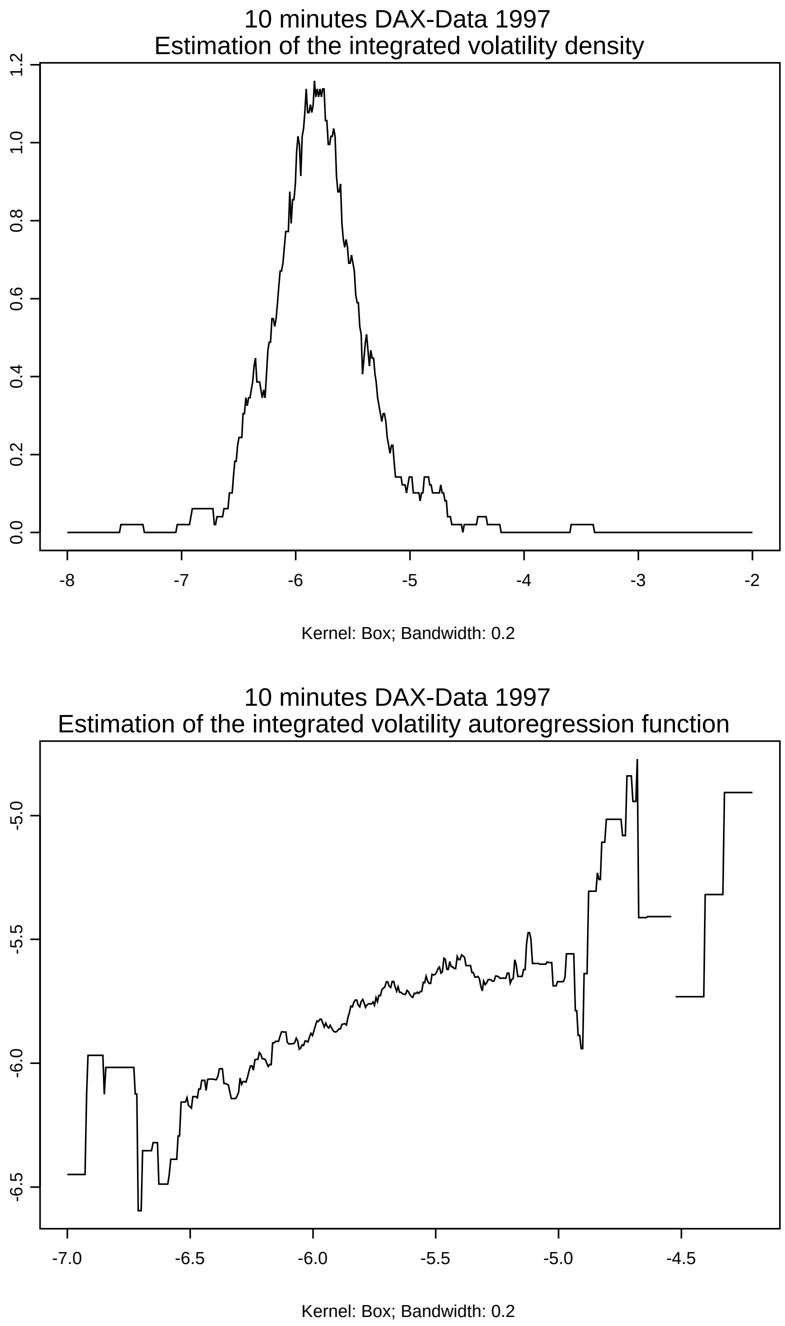

10 minutes DAX-Data 1997
Estimation of the integrated volatility density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
-8
-7
-6
-5
-4
-3
-2
Kernel: Box; Bandwidth: 0.2
10 minutes DAX-Data 1997
Estimation of the integrated volatility autoregression function
-6.5
-6.0
-5.5
-5.0
-7.0
-6.5
-6.0
-5.5
-5.0
-4.5
Kernel: Box; Bandwidth: 0.2

Figure 6.5:

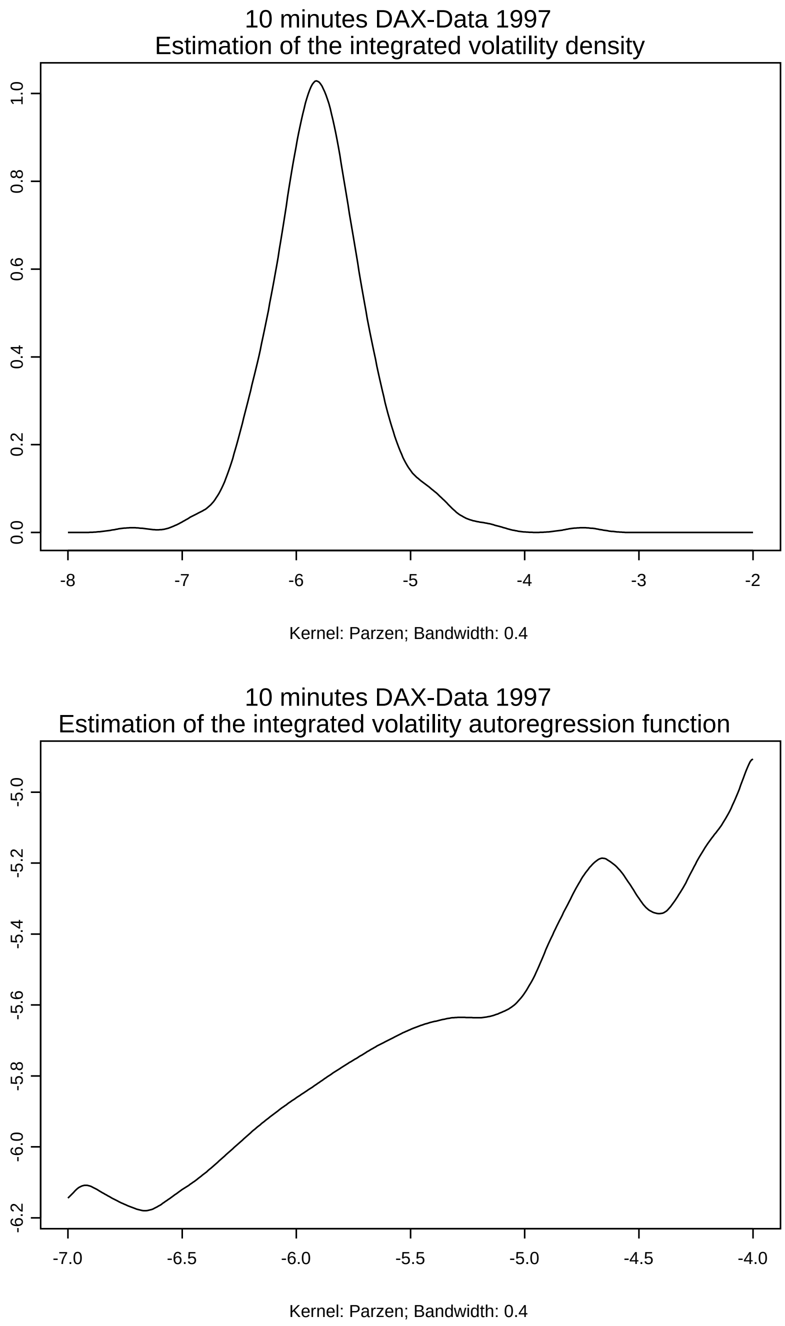

10 minutes DAX-Data 1997
Estimation of the integrated volatility density
0.0
0.2
0.4
0.6
0.8
1.0
-8
-7
-6
-5
-4
-3
-2
Kernel: Parzen; Bandwidth: 0.4
10 minutes DAX-Data 1997
Estimation of the integrated volatility autoregression function
-6.2
-6.0
-5.8
-5.6
-5.4
-5.2
-5.0
-7.0
-6.5
-6.0
-5.5
-5.0
-4.5
-4.0
Kernel: Parzen; Bandwidth: 0.4

Figure 6.6:

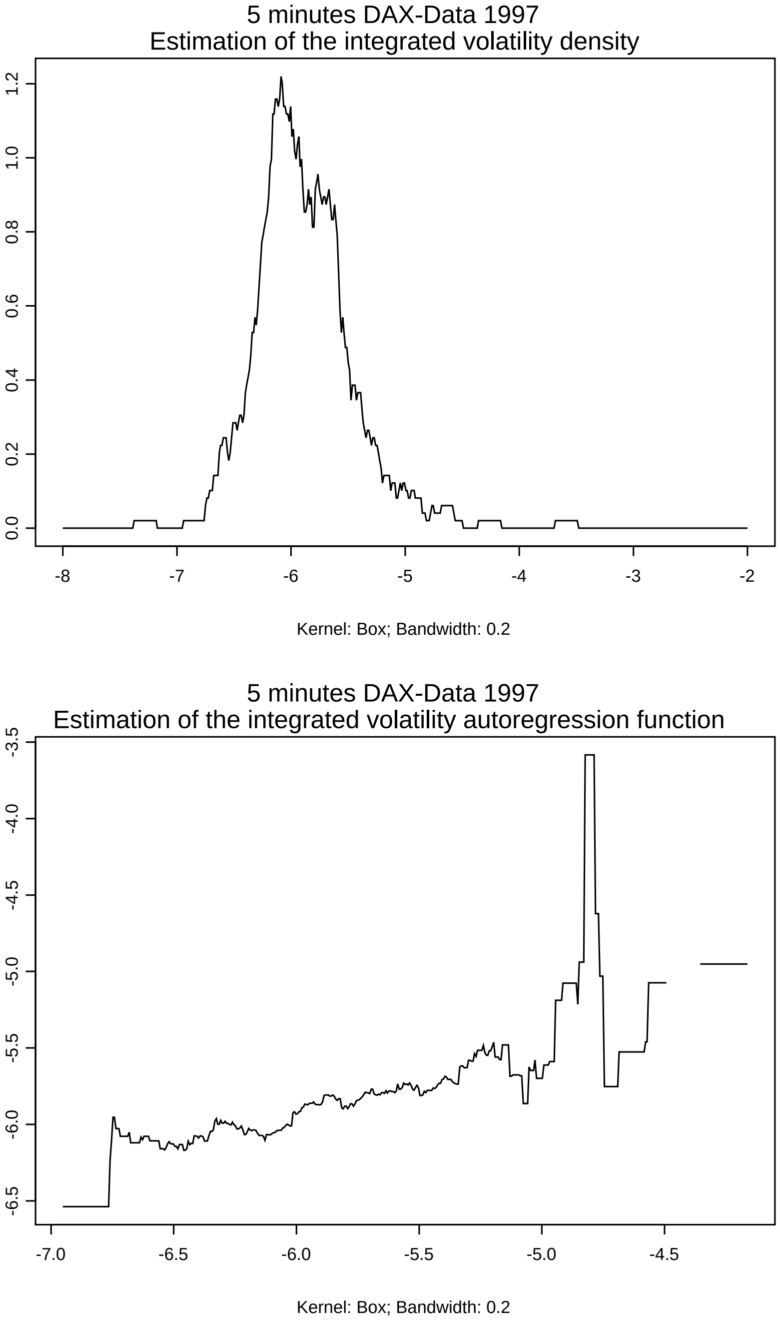

5 minutes DAX-Data 1997
Estimation of the integrated volatility density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
-8
-7
-6
-5
-4
-3
-2
Kernel: Box; Bandwidth: 0.2
5 minutes DAX-Data 1997
Estimation of the integrated volatility autoregression function
-6.5
-6.0
-5.5
-5.0
-4.5
-4.0
-3.5
-7.0
-6.5
-6.0
-5.5
-5.0
-4.5
Kernel: Box; Bandwidth: 0.2

Figure 6.7:

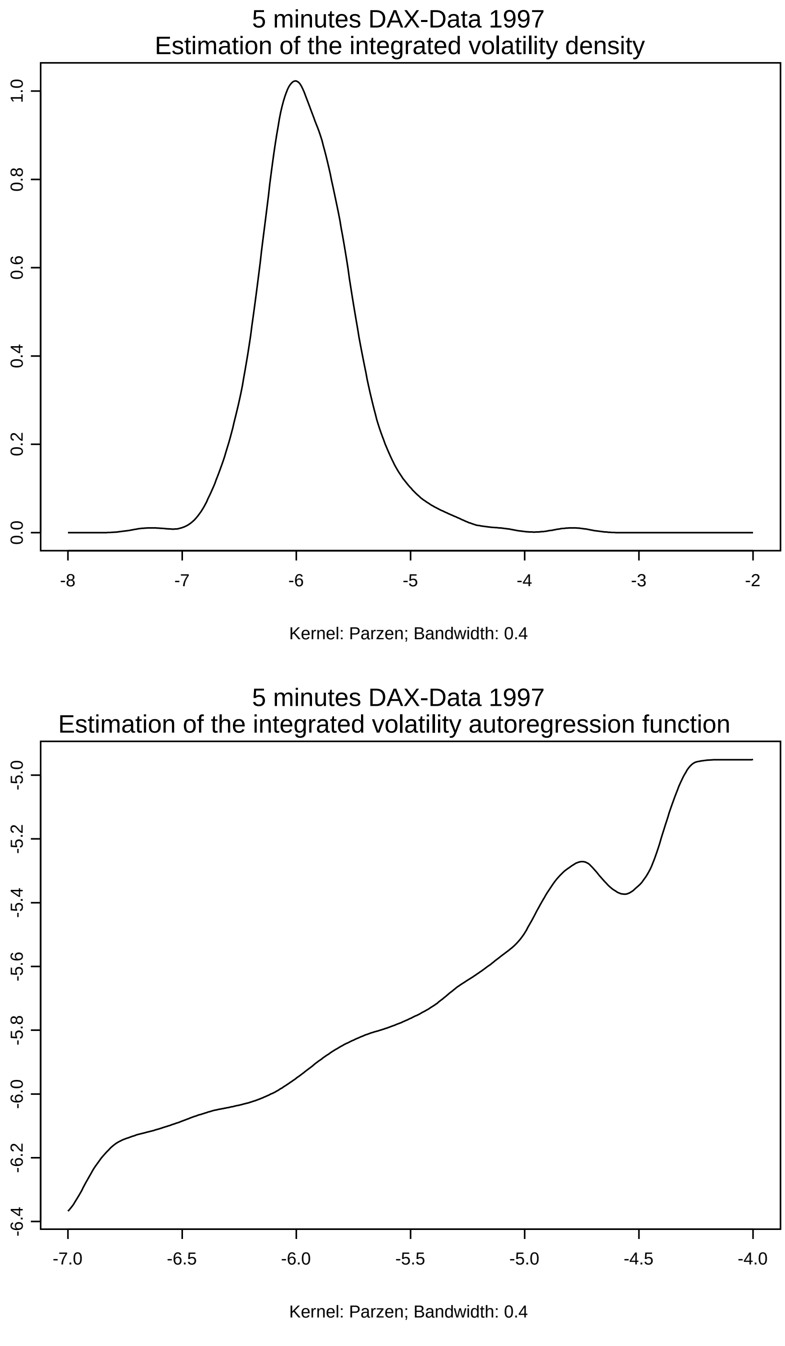

5 minutes DAX-Data 1997
Estimation of the integrated volatility density
0.0
0.2
0.4
0.6
0.8
1.0
-8
-7
-6
-5
-4
-3
-2
Kernel: Parzen; Bandwidth: 0.4
5 minutes DAX-Data 1997
Estimation of the integrated volatility autoregression function
-6.4
-6.2
-6.0
-5.8
-5.6
-5.4
-5.2
-5.0
-7.0
-6.5
-6.0
-5.5
-5.0
-4.5
-4.0
Kernel: Parzen; Bandwidth: 0.4

Figure 6.8:

# Chapter 7

# Estimation in a nonparametric GARCH(1,1) Model

As already mentioned in the introduction a shortcoming of standard ARCH and GARCH models is not to cope with the so called leverage effect. This is the reason, why we want to consider a nonparametric GARCH(1,1)-model, which is not affected by the symmetry-dilemma. Notice, that the proofs of all the results presented in this chapter are deferred to Appendix C

**7.1 Definition** (Nonparametric GARCH(1,1)).
*We define a nonparametric GARCH(1,1) process* $(X_t)_{t\in\mathbb{Z}}$ *as a process satisfying*

$$X_t = \sigma_t e_t \tag{7.1}$$

*with i.i.d. random-variables* $e_t$, $t \in \mathbb{Z}$, *with* $\mathrm{E}(e_t) = 0$, $\mathrm{E}(e_t^2) = \sigma_e^2$ *and* $\sigma_t$, $t \in \mathbb{Z}$, *following the nonparametric structure*

$$\sigma_t^2 = m\left(X_{t-1}, \sigma_{t-1}^2\right) = m\left(\sigma_{t-1} e_{t-1}, \sigma_{t-1}^2\right) \tag{7.2}$$

*with*

$$m(x, y) = m_1(x^2, y) 1_{(-\infty,0)}(x) + m_2(x^2, y) 1_{[0,\infty)}(x), \tag{7.3}$$

*where both* $m_i$*'s are Lipschitz-continous in both components, with*

$$|m_i(x, y) - m_i(x', y')| \leq L_{1,i} |x - x'| + L_{2,i} |y - y'|, \qquad i = 1, 2 \tag{7.4}$$

*and*

$$L := \max(L_{1,1}\sigma_e^2 + L_{2,1}, L_{1,2}\sigma_e^2 + L_{2,2}) \tag{7.5}$$

**7.2 Lemma** (Existence and Stationarity of the Nonparametric GARCH(1,1) process).
*For given i.i.d. random-variables* $e_t$, $t \in \mathbb{Z}$, *with* $\mathrm{E}\,(e_t) = 0$, $\mathrm{E}\,(e_t^2) = \sigma_e^2$ *equation* (7.2) *admits an unique stationary ergodic solution* $(\sigma_t)_{t\in\mathbb{Z}}$ *if* $L < 1$.
*Thus, if* $L < 1$, $X_t = \sigma_t e_t$, $t \in \mathbb{Z}$, *is also stationary.*

If one chooses $m_i(x,y) = \alpha_0 + \alpha_1 x^2 + \beta_1 y$, $i = 1,2$, $L < 1$ means $\alpha_1 + \beta_1 < 1$, which is the well known condition to ensure the existence of a GARCH(1,1)-process with existing second moments. We will show that - under the assumption $L < 1$ - the nonparametric GARCH(1,1) process is a weak-dependent stationary stochastic process. The concept of weak-dependence, which makes explicit a certain asymptotic independence, has been introduced by Doukhan and Louhichi in 1999 [35] and afterwards been investigated for example in Coulon-Prieur and Doukhan (2000) [28] and Ango Nze, Bühlmann and Doukhan (2002) [6].

**7.3 Definition** (A class $\mathcal{X}$ of weakly dependent real-valued, stationary stochastic processes).
*A real-valued stationary stochastic process* $(X_t)_t$ *belongs to the class* $\mathcal{X}$ *if there exists a constant* $\rho_X \in [0,1)$, *such that for any u-tuple* $(t_1,\ldots,t_u)$ *and any v-tuple* $(s_1,\ldots,s_v)$ *with* $t_1 \leq \ldots \leq t_u < s_1 \ldots \leq s_v$ *and arbitrary functions* $g : \mathbb{R}^u \to \mathbb{R}$, $h : \mathbb{R}^v \to \mathbb{R}$ *with* $\mathrm{E}\left(g\left(X_{t_1},\ldots,X_{t_u}\right)^2\right) < \infty$, $\mathrm{E}\left(h\left(X_{s_1},\ldots,X_{s_v}\right)^2\right) < \infty$ *there exists another constant* $C(u,v)$, *such that the following inequality is fulfilled:*

$$\begin{aligned} &\mathrm{Cov}\left(g\left(X_{t_1},\ldots,X_{t_u}\right), h\left(X_{s_1},\ldots,X_{s_v}\right)\right) \\ &\qquad \leq \sqrt{\mathrm{E}\left(g\left(X_{t_1},\ldots,X_{t_u}\right)^2\right)}\,\mathrm{Lip}\,(h)\, C(u,v)\rho_X^{s_1-t_u}, \end{aligned} \tag{7.6}$$

*where for any function* $h : \mathbb{R}^u \to \mathbb{R}$, $\mathrm{Lip}\,(h)$ *denotes its Lipschitz modulus of continuity defined as*

$$\mathrm{Lip}\,(h) = \sup_{x\neq y} \frac{|h(x)-h(y)|}{\|x-y\|_{l_1}}$$

*with* $\|x\|_{l_1} = \sum_{i=1}^u |x_i|$.

**7.4 Lemma.**
*If* $L < 1$ *the nonparametric GARCH(1,1)-process belongs to* $\mathcal{X}$.

Suppose that we observe such a nonparametric GARCH(1,1) process $(X_t)_{t\in\mathbb{Z}}$. The density $f_X$ of the stationary distribution of this process, which is assumed to exist, can then be estimated via a usual Kernel-estimator:

$$\hat{f}_X^T(x) = \frac{1}{T\lambda_T}\sum_{t=1}^{T} K\left(\frac{x - X_t}{\lambda_T}\right), \tag{7.7}$$

where $\lambda_T$ denotes again the so-called bandwidth, and K a probability density. Let again $U$ denote a random variable following this probability density.
By making use of the weak dependence of the nonparametric GARCH(1,1) process it can be shown, that - under certain assumptions - this estimator is asymptotically normal, which is formulated in the following Theorem:

**7.5 Theorem.**
*If $(X_t)_{t\in\mathbb{Z}}$ is of nonparametric GARCH(1,1)-structure with $L < 1$, $\lambda_T = T^{-\delta}$ with $0.2 < \delta < 1$ and if $\left\|f^{(i)}\right\|_\infty < \infty$, $i = 0, 1, 2$, $\|K\|_\infty < \infty$, $\mathrm{E}\left(|U|\right) < \infty$, $\mathrm{E}\left(U^2\right) < \infty$, then $\hat{f}_X^T$, defined in (7.7), fulfills*

$$\sqrt{T\lambda_T}\left(f_X(x) - \hat{f}_X^T(x)\right) \xrightarrow[N\to\infty]{d} \mathcal{N}(0, f_X(x)\,\|K\|_{L_2}^2).$$

Suppose that we observe the closing price $(S_t)_{t=1,\ldots,T}$, $T \in I\!N$, of a certain asset and that the log-return

$$X_t := \log\left(S_t\right) - \log\left(S_{t-1}\right)$$

of such an asset follows the nonparametric GARCH(1,1)-structure given in definition 7.1. As described in chapter 4, we assume, that with increasing sample size $T$ we observe an increasing number of (homogenously distributed) prices of the asset during the day, denoted by $S_{t,i}^T$, $i = 0, \ldots, M_T$.
For $i = 1, \ldots, M_T$, we define as in (4.1) the intra-day log-returns

$$X_{t,i}^T := \log\left(S_{t,i}^T\right) - \log\left(S_{t,i-1}^T\right)$$

and assume, that

$$X_{t,i}^T = \sigma_t \kappa_{t,i}^T \eta_{t,i}^T,$$

where $\kappa_{t,i}^T$ and $\eta_{t,i}^T$ have to fulfill the following assumptions:

**7.6 Assumptions.**

**A17** *The random variables* $\eta_{i,j}^T$, $T \in I\!N, i \in \{1, \dots, T+1\}, j \in \{1, \dots, M_T\}$, *are i.i.d., centered, of variance 1 and*

**A18** $\kappa_{i,j}^T$, $T \in I\!N, i \in \{1, \dots, T+1\}, j \in \{1, \dots, M_T\}$, *are deterministic with*

$$\sum_{j=1}^{M_T} \left(\kappa_{i,j}^T\right)^2 = 1.$$

Supposable one has to extend these assumptions - such as existence of some moments of the errors - to ensure the needed asymptotics, so that $\sigma_t^2$ can be estimated by

$$\begin{aligned}
\hat{\sigma_t^2}^T &= \frac{1}{M_t} \sum_{i=1}^{M_T} \left(X_{t,i}^T\right)^2 = \sigma_t^2 \cdot \left( \frac{1}{M_t} \sum_{i=1}^{M_T} \left(\kappa_{t,i}^T\right)^2 \left(\eta_{t_i}^T\right)^2 \right) \\
&= \sigma_t^2 + \sigma_t^2 \left( \frac{1}{M_t} \sum_{i=1}^{M_T} \left(\kappa_{t,i}^T\right)^2 \left(\eta_{t_i}^T\right)^2 - 1 \right)
\end{aligned}$$

and one can estimate the joint distribution of $X$ and $\sigma^2$ by

$$\hat{f}_{(X,\sigma^2)}^T(x,y) = \frac{1}{T\lambda_T^2} \sum_{t=1}^{T} K\left(\frac{x - X_t}{\lambda_T}\right) K\left(\frac{y - \hat{\sigma_t^2}^T}{\lambda_T}\right)$$

and the autoregression function $m$ by

$$\hat{m}^T(x,y) = \frac{\frac{1}{T\lambda_T^2} \sum_{t=1}^{T} K\left(\frac{x-X_t}{\lambda_T}\right) K\left(\frac{y-\hat{\sigma_t^2}^T}{\lambda_T}\right) X_{t+1}}{\hat{f}_{(X,\sigma^2)}^T(x,y)}.$$

# Appendix A

# Proofs of Chapter 2

## A.1 Proof of Theorem 2.2

In order to prove Theorem 2.2 we need the following auxiliary results

**A.1 Lemma.**
*If A1 and A2 are fulfilled, then*

*a)*

$$\operatorname{Var}\left(\hat{\phi}_X^T(\tau)\right) \leq \frac{1}{T}\left(\frac{9-\rho_\xi}{1-\rho_\xi}\right)$$

*b)*

$$\operatorname{Var}\left(\hat{\phi}_Z^T(\tau)\right) \leq \frac{1}{T}$$

**Proof:**

*a)*

$$
\begin{aligned}
&\operatorname{Var}\left(\frac{1}{T\cdot M}\sum_{j=1}^{T}\sum_{m=1}^{M}e^{i\tau X_{j,m}}\right)\\
&\quad=\operatorname{Var}\left(\frac{1}{T\cdot M}\sum_{j=1}^{T}\sum_{m=1}^{M}e^{i\tau(\xi_j+\epsilon_{j,m})}\right)\\
&\quad=\frac{1}{T^2M^2}\operatorname{Cov}\left(\sum_{j=1}^{T}\sum_{m=1}^{M}e^{i\tau(\xi_j+\epsilon_{j,m})},\sum_{k=1}^{T}\sum_{l=1}^{M}e^{i\tau\left(\xi_k+\epsilon_{k,l}\right)}\right)\\
&\quad=\frac{1}{T^2M^2}\left[\sum_{j=1}^{T}\sum_{m=1}^{M}\operatorname{Var}\left(e^{i\tau(\xi_j+\epsilon_{j,m})}\right)\right.\\
&\qquad\qquad+\sum_{j=1}^{T}\sum_{m=1}^{M}\sum_{\substack{l=1\\ l\neq m}}^{M}\operatorname{Cov}\left(e^{i\tau(\xi_j+\epsilon_{j,m})},e^{i\tau\left(\xi_j+\epsilon_{j,l}\right)}\right)\\
&\qquad\qquad\left.+\sum_{j=1}^{T}\sum_{\substack{k=1\\ k\neq j}}^{T}\sum_{m,l=1}^{M}\operatorname{Cov}\left(e^{i\tau(\xi_j+\epsilon_{j,m})},e^{i\tau\left(\xi_k+\epsilon_{k,l}\right)}\right)\right],
\end{aligned}
\tag{A.1.1}
$$

*where for* $j\in\{1,\dots,T\}$ *and* $m\in\{1,\dots,M\}$

$$
\begin{aligned}
\operatorname{Var}\left(e^{i\tau X_{j,m}}\right)&=\operatorname{E}\left(e^{i\tau X_{j,m}}e^{-i\tau X_{j,m}}\right)-\operatorname{E}\left(e^{i\tau X_{j,m}}\right)\operatorname{E}\left(e^{-i\tau X_{j,m}}\right)\\
&=1-\left|\phi_X(\tau)\right|^2\leq 1,
\end{aligned}
\tag{A.1.2}
$$

*for* $j\in\{1,\dots,T\}$, $m,l\in\{1,\dots,M\}$ *and* $m\neq l$

$$
\begin{aligned}
&\left|\operatorname{Cov}\left(e^{i\tau(\xi_j+\epsilon_{j,m})},e^{i\tau\left(\xi_j+\epsilon_{j,l}\right)}\right)\right|\\
&\qquad=\left|\operatorname{E}\left(e^{i\tau(\xi_j+\epsilon_{j,m})}e^{-i\tau\left(\xi_j+\epsilon_{j,l}\right)}\right)\right.\\
&\qquad\qquad\left.-\operatorname{E}\left(e^{i\tau(\xi_j+\epsilon_{j,m})}\right)\operatorname{E}\left(e^{-i\tau\left(\xi_j+\epsilon_{j,l}\right)}\right)\right|\\
&\qquad=\left|\operatorname{E}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,l})}\right)-\phi_X(\tau)\phi_X(-\tau)\right|\\
&\qquad=\left|\phi_\epsilon^2(\tau)-\left|\phi_X(\tau)\right|^2\right|\leq 1,
\end{aligned}
\tag{A.1.3}
$$

*and for* $j, k \in \{1, \ldots, T\}$, $j \neq k$ *and* $m, l \in \{1, \ldots, M\}$

$$\begin{aligned}
&\left|\mathrm{Cov}\left(e^{i\tau(\xi_j+\epsilon_{j,m})}, e^{i\tau\left(\xi_k+\epsilon_{k,l}\right)}\right)\right| \\
&= \left|\mathrm{E}\left(e^{i\tau(\xi_j-\xi_k)}e^{i\tau\left(\epsilon_{j,m}-\epsilon_{k,l}\right)}\right) - \mathrm{E}\left(e^{i\tau\xi_j}e^{i\tau\epsilon_{j,m}}\right)\mathrm{E}\left(e^{-i\tau\xi_k}e^{-i\tau\epsilon_{k,l}}\right)\right| \\
&= \left|\mathrm{E}\left(e^{i\tau\xi_j}e^{-i\tau\xi_k}\right)\mathrm{E}\left(e^{i\tau\epsilon_{j,m}}\right)\mathrm{E}\left(e^{-i\tau\epsilon_{k,l}}\right)\right. \\
&\qquad \left. -\mathrm{E}\left(e^{i\tau\xi_j}\right)\mathrm{E}\left(e^{-i\tau\xi_k}\right)\mathrm{E}\left(e^{i\tau\epsilon_{j,m}}\right)\mathrm{E}\left(e^{-i\tau\epsilon_{k,l}}\right)\right| \\
&= \phi_\epsilon^2(\tau)\left|\mathrm{E}\left(e^{i\tau\xi_j}e^{-i\tau\xi_k}\right) - \mathrm{E}\left(e^{i\tau\xi_j}\right)\mathrm{E}\left(e^{-i\tau\xi_k}\right)\right| \\
&= \phi_\epsilon^2(\tau)\Big|\mathrm{E}\left([\cos(\tau\xi_j) + i\sin(\tau\xi_j)]\,[\cos(\tau\xi_k) - i\sin(\tau\xi_k)]\right) \\
&\qquad -\mathrm{E}\left(\cos(\tau\xi_j) + i\sin(\tau\xi_j)\right)\mathrm{E}\left(\cos(\tau\xi_k) - i\sin(\tau\xi_k)\right)\Big| \\
&= \phi_\epsilon^2(\tau)\Big|\mathrm{E}\left(\cos(\tau\xi_j)\cos(\tau\xi_k)\right) - i\mathrm{E}\left(\cos(\tau\xi_j)\sin(\tau\xi_k)\right) \\
&\qquad +i\mathrm{E}\left(\sin(\tau\xi_j)\cos(\tau\xi_k)\right) + \mathrm{E}\left(\sin(\tau\xi_j)\sin(\tau\xi_k)\right) \\
&\qquad -\Big(\mathrm{E}\left(\cos(\tau\xi_j)\right)\mathrm{E}\left(\cos(\tau\xi_k)\right) - i\mathrm{E}\left(\cos(\tau\xi_j)\right)\mathrm{E}\left(\sin(\tau\xi_k)\right) \\
&\qquad\quad +i\mathrm{E}\left(\sin(\tau\xi_j)\right)\mathrm{E}\left(\cos(\tau\xi_k)\right) + \mathrm{E}\left(\sin(\tau\xi_j)\right)\mathrm{E}\left(\sin(\tau\xi_k)\right)\Big)\Big| \\
&= \phi_\epsilon^2(\tau)\Big|\mathrm{Cov}\left(\cos(\tau\xi_j), \cos(\tau\xi_k)\right) - i\mathrm{Cov}\left(\cos(\tau\xi_j), \sin(\tau\xi_k)\right) \\
&\qquad +i\mathrm{Cov}\left(\sin(\tau\xi_j), \cos(\tau\xi_k)\right) + \mathrm{Cov}\left(\sin(\tau\xi_j), \sin(\tau\xi_k)\right)\Big| \\
&\leq \phi_\epsilon^2(\tau)4\alpha_\xi(|j-k|) \leq 4\rho_\xi^{|j-k|},
\end{aligned} \tag{A.1.4}$$

*where the two last inequalities are due to the fact, that* $(\xi_j)_{j\in\mathbb{N}}$ *is* $\alpha$*-mixing and to* (1.8)*. Plugging the results of the equationarrays* (A.1.2) *to* (A.1.4) *into equation* (A.1.1) *we achieve:*

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{T\cdot M}\sum_{j=1}^{T}\sum_{m=1}^{M} e^{i\tau X_{j,m}}\right) \\
&\leq \frac{1}{T^2M^2}\left(TM + TM(M-1) + 2\sum_{j=1}^{T}\sum_{k=i+1}^{T} M^2 4\rho_\xi^{k-j}\right) \\
&\leq \frac{1}{T^2M^2}\left(TM^2 + 8M^2\sum_{i=1}^{T}\frac{1}{1-\rho_\xi}\right) = \frac{1}{T^2M^2}TM^2\left(1 + \frac{8}{1-\rho_\xi}\right) \\
&= \frac{1}{T}\left(\frac{9-\rho_\xi}{1-\rho_\xi}\right),
\end{aligned}$$

*b) The proof is almost the same as the one of part a):*

$$
\begin{aligned}
&\operatorname{Var}\left(\frac{1}{T(M-1)}\sum_{j=1}^{T}\sum_{m=2}^{M} e^{i\tau Z_{j,m}}\right)\\
&= \frac{1}{T^2(M-1)^2}\operatorname{Cov}\left(\sum_{j=1}^{T}\sum_{m=2}^{M} e^{i\tau(\epsilon_{j,m}-\epsilon_{j,1})}, \sum_{k=1}^{T}\sum_{l=2}^{M} e^{i\tau(\epsilon_{k,l}-\epsilon_{k,1})}\right)\\
&= \frac{1}{T^2(M-1)^2}\left[\sum_{j=1}^{T}\sum_{m=2}^{M}\operatorname{Var}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,1})}\right)\right. \qquad (A.1.5)\\
&\quad + \sum_{j=1}^{T}\sum_{m=2}^{M}\sum_{\substack{l=2\\ l\neq m}}^{M}\operatorname{Cov}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,1})}, e^{i\tau(\epsilon_{j,l}-\epsilon_{j,1})}\right)\\
&\quad \left. + \sum_{j=1}^{T}\sum_{\substack{k=1\\ k\neq j}}^{T}\sum_{m,l=2}^{M}\operatorname{Cov}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,1})}, e^{i\tau(\epsilon_{k,l}-\epsilon_{k,1})}\right)\right],
\end{aligned}
$$

*where for $j \in \{1,\ldots,T\}$ and $m \in \{2,\ldots,M\}$*

$$
\begin{aligned}
&\left|\operatorname{Var}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,1})}\right)\right|\\
&\quad = \left|\operatorname{E}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,1})}e^{i\tau(\epsilon_{j,1}-\epsilon_{j,m})}\right) - \phi_\epsilon(\tau)\phi_\epsilon(-\tau)\phi_\epsilon(\tau)\phi_\epsilon(-\tau)\right| \qquad (A.1.6)\\
&\quad = \left|1-\phi_\epsilon^4(\tau)\right| \leq 1
\end{aligned}
$$

*for $j \in \{1,\ldots,T\}$; $m,l \in \{1,\ldots,M\}$ and $m \neq l$:*

$$
\begin{aligned}
&\left|\operatorname{Cov}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,1})}, e^{i\tau(\epsilon_{j,l}-\epsilon_{j,1})}\right)\right|\\
&\qquad = \left|\operatorname{E}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,l})}\right) - \operatorname{E}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,1})}\right)\operatorname{E}\left(e^{i\tau(\epsilon_{j,1}-\epsilon_{j,l})}\right)\right|\\
&\qquad = \left|\phi_\epsilon(\tau)\phi_\epsilon(-\tau) - \phi_\epsilon(\tau)\phi_\epsilon(-\tau)\phi_\epsilon(\tau)\phi_\epsilon(-\tau)\right| \qquad (A.1.7)\\
&\qquad = \phi_\epsilon^2(\tau)\left|1-\phi_\epsilon^2(\tau)\right|\\
&\qquad \leq 1,
\end{aligned}
$$

*and for $j,k \in \{1,\ldots,T\}$; $j \neq k$ and $m,l \in \{2,\ldots,M\}$*

$$
\operatorname{Cov}\left(e^{i\tau(\epsilon_{j,m}-\epsilon_{j,1})}, e^{i\tau(\epsilon_{k,l}-\epsilon_{k,1})}\right) = 0. \qquad (A.1.8)
$$

*Plugging the results of the equationarrays* (A.1.6) *to* (A.1.8) *into equation* (A.1.5) *we achieve:*

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{1}{T(M-1)}\sum_{j=1}^{T}\sum_{m=2}^{M} e^{i\tau Z_{j,m}}\right)\\
&\quad\leq \frac{1}{T^2(M-1)^2}\left(T(M-1)+2T(M-1)(M-2)\right)\\
&\quad= \frac{1}{T^2(M-1)^2}T(M-1)(1+M-2)\\
&\quad= \frac{1}{T}.
\end{aligned}
$$

□

**A.2 Lemma.**
*If A1 and A2 are fulfilled, then*

*a)*

$$\mathrm{P}\left(\left|\hat{\phi}_X^T(\tau)-\phi_X(\tau)\right|>\epsilon\right)\leq\frac{1}{T\epsilon^2}\left(\frac{3-\rho_\xi}{9-\rho_\xi}\right)$$

*b)*

$$\mathrm{P}\left(\left|\hat{\phi}_Z^T(\tau)-\phi_Z(\tau)\right|>\epsilon\right)\leq\frac{1}{T\epsilon^2}$$

**Proof:**
*It is obvious, that,* $\mathrm{E}\left(\hat{\phi}_X^T(\tau)\right)=\phi_X(\tau)$ *and* $\mathrm{E}\left(\hat{\phi}_Z^T(\tau)\right)=\phi_Z(\tau)$*. Easy Computation shows, that Chebyshev's inequality is valid even in this complex case and by using Lemma A.1a) and b), we get:*

$$\mathrm{P}\left(\left|\hat{\phi}_X^T(\tau)-\phi_X(\tau)\right|>\epsilon\right)\leq\frac{\mathrm{Var}\left(\hat{\phi}_X^T(\tau)\right)}{\epsilon^2}=\frac{1}{T\epsilon^2}\left(\frac{3-\rho_\xi}{9-\rho_\xi}\right)$$

*and*

$$\mathrm{P}\left(\left|\hat{\phi}_Z^T(\tau)-\phi_Z(\tau)\right|>\epsilon\right)\leq\frac{\mathrm{Var}\left(\hat{\phi}_Z^T(\tau)\right)}{\epsilon^2}=\frac{1}{T\epsilon^2}.$$

□

**A.3 Lemma.**
*If A1, A2, A5 are fulfilled and* $\tau \in \left[-\frac{1}{\lambda_T}, \frac{1}{\lambda_T}\right]$, *then*

$$\mathrm{P}\left(\left|\tilde{\phi}_Z^T(\tau) - \phi_Z(\tau)\right| > \epsilon\right) \leq \frac{1}{T\epsilon^2} + \frac{4}{Tc_T^4}.$$

**Proof:**

$$\begin{aligned}
&\mathrm{P}\left(\left|\tilde{\phi}_Z^T(\tau) - \phi_Z(\tau)\right| > \epsilon\right) \\
\leq\;& \mathrm{P}\left(\left|\tilde{\phi}_Z^T(\tau) - \phi_Z(\tau)\right| > \epsilon, \mathcal{R}e\left(\hat{\phi}_Z^T(\tau)\right) \geq \frac{c_T^2}{2}\right) + \mathrm{P}\left(\mathcal{R}e\left(\hat{\phi}_Z^T(\tau)\right) < \frac{c_T^2}{2}\right) \\
\leq\;& \mathrm{P}\left(\left|\mathcal{R}e\left(\hat{\phi}_Z^T(\tau)\right) - \mathcal{R}e\left(\phi_Z(\tau)\right)\right| > \epsilon, \mathcal{R}e\left(\hat{\phi}_Z^T(\tau)\right) \geq \frac{c_T^2}{2}\right) \\
&+\mathrm{P}\left(\left|\mathcal{R}e\left(\hat{\phi}_Z^T(\tau)\right) - \mathcal{R}e\left(\phi_Z(\tau)\right)\right| > \frac{c_T^2}{2}\right) \\
\leq\;& \mathrm{P}\left(\left|\mathcal{R}e\left(\hat{\phi}_Z^T(\tau)\right) - \mathcal{R}e\left(\phi_Z(\tau)\right)\right| > \epsilon\right) + \mathrm{P}\left(\left|\mathcal{R}e\left(\hat{\phi}_Z^T(\tau)\right) - \mathcal{R}e\left(\phi_Z(\tau)\right)\right| > \frac{c_T^2}{2}\right) \\
\leq\;& \mathrm{P}\left(\left|\hat{\phi}_Z^T(\tau) - \phi_Z(\tau)\right| > \epsilon\right) + \mathrm{P}\left(\left|\hat{\phi}_Z^T(\tau) - \phi_Z(\tau)\right| > \frac{c_T^2}{2}\right) \\
\leq\;& \frac{1}{T\epsilon^2} + \frac{4}{Tc_T^4},
\end{aligned}$$

*where the second inequality is due to the fact, that* $\phi_Z(\tau) > c_T^2$ *for* $\tau \in \left[-\frac{1}{\lambda_T}, \frac{1}{\lambda_T}\right]$ *and the last inequality is due to Lemma A.2b).*

□

**A.4 Lemma.**
*If A1, A2, A5 are fulfilled and* $\tau \in \left[-\frac{1}{\lambda_T}, \frac{1}{\lambda_T}\right]$, *then*

$$\mathrm{P}\left(\left|\hat{\phi}_\epsilon^T(\tau) - \phi_\epsilon(\tau)\right| > \epsilon\right) \leq \frac{1}{Tc_T^2\epsilon^2} + \frac{4}{Tc_T^4}.$$

**Proof:**
*Recall that* $\phi_\epsilon > c_T$ *on* $\left[-\frac{1}{\lambda_T}, \frac{1}{\lambda_T}\right]$. *Thus we have*

$$\left|\hat{\phi}_\epsilon^T(\tau) - \phi_\epsilon(\tau)\right| = \frac{\left|\tilde{\phi}_Z^T(\tau) - \phi_Z(\tau)\right|}{\left|\hat{\phi}_\epsilon^T(\tau) + \phi_\epsilon(\tau)\right|} \leq \frac{\left|\tilde{\phi}_Z^T(\tau) - \phi_Z(\tau)\right|}{\phi_\epsilon(\tau)} \leq \frac{1}{c_T}\left|\tilde{\phi}_Z^T(\tau) - \phi_Z(\tau)\right|,$$

*which implies*

$$
\begin{aligned}
\mathrm{P}\left(\left|\hat{\phi}_\epsilon^T(\tau) - \phi_\epsilon(\tau)\right| > \epsilon\right) &\leq \mathrm{P}\left(\frac{1}{c_T}\left|\tilde{\phi}_Z^T(\tau) - \phi_Z(\tau)\right| > \epsilon\right) \\
&= \mathrm{P}\left(\left|\tilde{\phi}_Z^T(\tau) - \phi_Z(\tau)\right| > c_T\epsilon\right) \\
&\leq \frac{1}{Tc_T^2\epsilon^2} + \frac{4}{Tc_T^4},
\end{aligned}
$$

*where the last inequality is due to Lemma A.3.*

□

To prove Theorem 2.2 we divide $\hat{f}_\xi^T(x) - f_\xi(x)$ into

$$
\hat{f}_\xi^T(x) - f_\xi(x) = \left(\hat{f}_\xi^T(x) - R_T(x)\right) + \left(R_T(x) - I_T(x)\right) + \left(I_T(x) - f_\xi(x)\right), \quad \text{(A.1)}
$$

where

$$
R_T(x) := \frac{1}{2\pi}\sum_{j=1}^{S_T} e^{-i\tau_j^T x}\phi_U(\lambda_T\tau_j)\frac{\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\,d_T, \quad \text{(A.2)}
$$

$$
I_T(x) := \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau x}\phi_U(\lambda_T\tau)\frac{\phi_X(\tau)}{\phi_\epsilon(\tau)}\,d\tau, \quad \text{(A.3)}
$$

with $d_T, S_T$ and $\tau_j^T$ defined as in (2.17), and study the limiting behaviour of the three summands separately.

**A.5 Lemma.**
*If A1, A2, A5, A6 and A7 are fulfilled, then*

*a)*

$$
\mathrm{P}\left(\left|\hat{f}_\xi^T(x) - R_T(x)\right| > \epsilon\right) \leq \frac{8(2+c_T^2)}{T\lambda_T d_T c_T^4} + \frac{2}{T\lambda_T^3 d_T^3 c_T^2\pi^2\epsilon^2}\left(\frac{10-2\rho_\xi}{1-\rho_\xi}\right),
$$

*b)*

$$
|R_T(x) - I_T(x)| \leq \frac{\mathrm{E}\left(|U|\right) + \mathrm{E}\left(|\xi|\right)}{2\pi}\cdot a_T,
$$

*c)*

$$
|I_T(x) - f_\xi(x)| \leq \lambda_T\left\|f_\xi'\right\|_\infty E(|U|).
$$

To prove Lemma A.5a) we will need the following Lemma:

**A.6 Lemma.**
*Let $S_T$ and $\tau_j^T$ be as defined in* (2.17). *If A1, A2 and A5 are fulfilled, then*

*a)*

$$\mathrm{P}\left(\max_{j\in\{1,\dots,S_T\}}\left|\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right|>\epsilon\right)\leq\frac{2}{T\lambda_T d_T c_T^2\epsilon^2}+\frac{8}{T\lambda_T d_T c_T^4},$$

*b)*

$$\mathrm{P}\left(\max_{j\in\{1,\dots,S_T\}}\left|\frac{\hat{\phi}_X^T(\tau_j^T)-\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right|>\epsilon\right)\leq\frac{2}{T\lambda_T d_T c_T^2\epsilon^2}\left(\frac{9-\rho_\xi}{1-\rho_\xi}\right).$$

**Proof:**
*Recall that by Assumption A5 $\phi_\epsilon(\tau)>c_T$. So we get for $\tau\in[-\lambda_T^{-1},\lambda_T^{-1}]$:*

$$\begin{aligned}
\mathrm{P}\left(\max_{j\in\{1,\dots,S_T\}}\left|\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right|>\epsilon\right) &= \mathrm{P}\left(\bigcup_{j=1}^{S_T}\left\{\left|\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right|>\epsilon\right\}\right)\\
&\leq \sum_{j=1}^{S_T}\mathrm{P}\left(\left|\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right|>\epsilon\right)\\
&= \sum_{j=1}^{S_T}\mathrm{P}\left(\left|\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)\right|>\epsilon\cdot\phi_\epsilon(\tau_j^T)\right)\\
&\leq \sum_{j=1}^{S_T}\mathrm{P}\left(\left|\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)\right|>\epsilon\cdot c_T\right)\\
&\leq S_T\left(\frac{1}{Tc_T^2\epsilon^2}+\frac{4}{Tc_T^4}\right)\\
&= \frac{2}{T\lambda_T d_T c_T^2\epsilon^2}+\frac{8}{T\lambda_T d_T c_T^4},
\end{aligned}$$

*where the last inequality is due to Lemma A.4.*
*The proof of part b) is almost the same, you just have to change the nominators and make use of Lemma A.2a) instead of Lemma A.4 in the last but one step.*

□

**Proof of Lemma A.5:**

a)

$$
\begin{aligned}
&\left|\hat{f}_\xi^T(x) - R_T(x)\right| \\
&\quad \leq \frac{S_T}{2\pi} \cdot \max_{j\in\{1,\dots,S_T\}} \left| \frac{\hat{\phi}_X^T(\tau_j^T)}{\hat{\phi}_\epsilon^T(\tau_j^T)} - \frac{\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} \right| \\
&\quad = \frac{S_T}{2\pi} \cdot \max_{j\in\{1,\dots,S_T\}} \left| \frac{\hat{\phi}_X^T(\tau_j^T)}{\hat{\phi}_\epsilon^T(\tau_j^T)} - \phi_\xi(\tau_j^T) \right| \\
&\quad = \frac{S_T}{2\pi} \cdot \max_{j\in\{1,\dots,S_T\}} \left| \frac{\hat{\phi}_X^T(\tau_j^T)}{\hat{\phi}_\epsilon^T(\tau_j^T)} - \frac{\phi_X(\tau_j^T)}{\hat{\phi}_\epsilon^T(\tau_j^T)} + \frac{\phi_\xi(\tau_j^T)\phi_\epsilon(\tau_j^T)}{\hat{\phi}_\epsilon^T(\tau_j^T)} - \frac{\phi_\xi(\tau_j^T)\hat{\phi}_\epsilon^T(\tau_j^T)}{\hat{\phi}_\epsilon^T(\tau_j^T)} \right| \\
&\quad = \frac{S_T}{2\pi} \cdot \max_{j\in\{1,\dots,S_T\}} \left| \frac{\frac{\hat{\phi}_X^T(\tau_j^T)-\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - \phi_\xi(\tau_j^T)\left(\frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1\right)}{1 + \left(\frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1\right)} \right|.
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\mathrm{P}\left(\left|\hat{f}_\xi^T(x) - R_T(x)\right| > \epsilon\right) \\
&\quad \leq \mathrm{P}\left( \max_{j\in\{1,\dots,S_T\}} \left| \frac{\frac{\hat{\phi}_X^T(\tau_j^T)-\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - \phi_\xi(\tau_j^T)\left(\frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1\right)}{1 + \left(\frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1\right)} \right| > \frac{2\pi\epsilon}{S_T} \right) \\
&\quad \leq \mathrm{P}\left( \frac{\max\limits_{j\in\{1,\dots,S_T\}} \left| \frac{\hat{\phi}_X^T(\tau_j^T)-\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} \right| + \max\limits_{j\in\{1,\dots,S_T\}} \left( \phi_\xi(\tau_j^T) \left| \frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1 \right| \right)}{\min\limits_{j\in\{1,\dots,S_T\}} \left( 1 + \left( \frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1 \right) \right)} > \pi\lambda_T d_T \epsilon \right) \\
&\quad \leq \mathrm{P}\left( \frac{\max\limits_{j\in\{1,\dots,S_T\}} \left| \frac{\hat{\phi}_X^T(\tau_j^T)-\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} \right| + \max\limits_{j\in\{1,\dots,S_T\}} \left( \phi_\xi(\tau_j^T) \left| \frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1 \right| \right)}{\min\limits_{j\in\{1,\dots,S_T\}} \left( 1 + \left( \frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1 \right) \right)} > \pi\lambda_T d_T \epsilon, \right. \\
&\qquad\qquad \left. \min_{j\in\{1,\dots,S_T\}} \left( 1 + \left( \frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1 \right) \right) \geq \frac{1}{2} \right) \\
&\qquad\quad + \mathrm{P}\left( \min_{j\in\{1,\dots,S_T\}} \left( 1 + \left( \frac{\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} - 1 \right) \right) < \frac{1}{2} \right) \\
&\quad \leq \mathrm{P}\left( \max_{j\in\{1,\dots,S_T\}} \left| \frac{\hat{\phi}_X^T(\tau_j^T) - \phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} \right| \right.
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \max_{j\in\{1,\dots,S_T\}}\left(\phi_\xi(\tau_j^T)\left|\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right|\right) > 2\pi\lambda_T d_T\epsilon\Bigg) \\
&\quad +\mathrm{P}\left(\min_{j\in\{1,\dots,S_T\}}\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} < -\frac{1}{2}\right) \\
\leq\ & \mathrm{P}\Bigg(\left\{\max_{j\in\{1,\dots,S_T\}}\left|\frac{\hat{\phi}_X^T(\tau_j^T)-\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right| > \pi\lambda_T d_T\epsilon\right\} \\
&\qquad \cup\left\{\max_{j\in\{1,\dots,S_T\}}\left(\phi_\xi(\tau_j^T)\left|\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right|\right) > \pi\lambda_T d_T\epsilon\right\}\Bigg) \\
&\quad +\mathrm{P}\left(\min_{j\in\{1,\dots,S_T\}}\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} < -\frac{1}{2}\right) \\
\leq\ & \mathrm{P}\left(\max_{j\in\{1,\dots,S_T\}}\frac{\phi_\epsilon(\tau_j^T)-\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)} > \frac{1}{2}\right) \\
&\quad +\mathrm{P}\left(\max_{j\in\{1,\dots,S_T\}}\left|\frac{\hat{\phi}_X^T(\tau_j^T)-\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right| > \pi\lambda_T d_T\epsilon\right) \\
&\quad +\mathrm{P}\left(\max_{j\in\{1,\dots,S_T\}}\left(\phi_\xi(\tau_j^T)\left|\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right|\right) > \pi\lambda_T d_T\epsilon\right) \\
\leq\ & \mathrm{P}\left(\max_{j\in\{1,\dots,S_T\}}\left|\frac{\phi_\epsilon(\tau_j^T)-\hat{\phi}_\epsilon^T(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right| > \frac{1}{2}\right) \\
&\quad +\mathrm{P}\left(\max_{j\in\{1,\dots,S_T\}}\left|\frac{\hat{\phi}_X^T(\tau_j^T)-\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right| > \pi\lambda_T d_T\epsilon\right) \\
&\quad +\mathrm{P}\left(\max_{j\in\{1,\dots,S_T\}}\left|\frac{\hat{\phi}_\epsilon^T(\tau_j^T)-\phi_\epsilon(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\right| > \pi\lambda_T d_T\epsilon\right) \\
\leq\ & \frac{8}{T\lambda_T d_T c_T^2}+\frac{8}{T\lambda_T d_T c_T^4}+\frac{2}{T\lambda_T d_T c_T^2\pi^2\lambda_T^2 d_T^2\epsilon^2}\left(\frac{9-\rho_\xi}{1-\rho_\xi}\right) \\
&\quad +\frac{2}{T\lambda_T d_T c_T^2\pi^2\lambda_T^2 d_T^2\epsilon^2}+\frac{8}{T\lambda_T d_T c_T^4} \\
=\ & \frac{8(2+c_T^2)}{T\lambda_T d_T c_T^4}+\frac{2}{T\lambda_T^3 d_T^3 c_T^2\pi^2\epsilon^2}\left(\frac{9-\rho_\xi}{1-\rho_\xi}+1\right) \\
=\ & \frac{8(2+c_T^2)}{T\lambda_T d_T c_T^4}+\frac{2}{T\lambda_T^3 d_T^3 c_T^2\pi^2\epsilon^2}\left(\frac{10-2\rho_\xi}{1-\rho_\xi}\right),
\end{aligned}
$$

where the last inequality is due to Lemma A.6a) and b).

b)

$$
\begin{aligned}
|R_T(x) - I_T(x)| &\leq \frac{1}{2\pi}\sum_{j=1}^{S_T}\Bigg| e^{-i\tau_j x}\phi_U(\lambda_T\tau_j)\frac{\phi_X(\tau_j^T)}{\phi_\epsilon(\tau_j^T)}\,d_T \\
&\qquad - \int_{\tau_j^T}^{\tau_{j+1}^T} e^{-i\tau x}\phi_U(\lambda_T\tau)\frac{\phi_X(\tau)}{\phi_\epsilon(\tau)}\,d\tau\Bigg| \\
&\leq \frac{1}{2\pi}\sum_{j=1}^{S_T}\frac{d_T^2}{2}\sup_{\tau\in[\tau_j^T,\tau_{j+1}^T]}\left|\frac{d}{d\tau}\left(\phi_U(\lambda_T\tau)\phi_\xi(\tau)\right)\right| \\
&= \frac{1}{2\pi}\sum_{j=1}^{S_T}\frac{d_T^2}{2}\sup_{\tau\in[\tau_j^T,\tau_{j+1}^T]}\left|\lambda_T\phi_U'(\lambda_T\tau)\phi_\xi(\tau)+\phi_U(\lambda_T\tau)\phi_\xi'(\tau)\right| \\
&\leq \frac{1}{2\pi}\sum_{j=1}^{S_T}\frac{d_T^2}{2}\left(\lambda_T\mathrm{E}\left(|U|\right)\cdot 1 + 1\cdot\mathrm{E}\left(|\xi|\right)\right) \\
&\leq \frac{\mathrm{E}\left(|U|\right)+\mathrm{E}\left(|\xi|\right)}{2\pi}\cdot S_T\cdot\frac{d_T^2}{2} = \frac{\mathrm{E}\left(|U|\right)+\mathrm{E}\left(|\xi|\right)}{2\pi}\cdot\frac{2}{\lambda_T d_T}\cdot\frac{d_t^2}{2} \\
&= \frac{\mathrm{E}\left(|U|\right)+\mathrm{E}\left(|\xi|\right)}{2\pi}\cdot\frac{d_t}{\lambda_T},
\end{aligned}
$$

where the second inequlity is just an application of Taylor's formula.

c)

$$
\begin{aligned}
I_T(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau x}\phi_U(\lambda_T\tau)\frac{\phi_X(\tau)}{\phi_\epsilon(\tau)}\,d\tau = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau x}\phi_{\xi+\lambda_T U}(\tau)\,d\tau \\
&= f_{\xi+\lambda_T U}(x) = \int_{-\infty}^{\infty} f_\xi(x-y)f_{\lambda_T U}(y)\,dy \\
&= \int_{-\infty}^{\infty} f_\xi(x-y)f_U(\lambda_T^{-1}y)\lambda_T^{-1}\,dy = \int_{-\infty}^{\infty} f_\xi(x-\lambda_T u)f_U(u)\,du \\
&= \int_{-\infty}^{\infty}\left(f_\xi(x)-\lambda_T u f_\xi'(\theta_u)\right)f_U(u)\,du \\
&= f_\xi(x) - \lambda_T\int_{-\infty}^{\infty} u f_\xi'(\theta_u)f_U(u)\,du.
\end{aligned}
$$

Thus

$$
|I_T(x) - f_\xi(x)| \leq \lambda_T\left\|f_\xi'\right\|_\infty\int_{-\infty}^{\infty}|u|\,f_U(u)\,du = \lambda_T\left\|f_\xi'\right\|_\infty E(|U|).
$$

□

**A.7 Lemma.**
*Let $\epsilon > 0$ and A1, A2, A5, A6 and A7 be fulfilled. If $T$ is chosen large enough, such that $a_T < \frac{\epsilon}{4}\frac{2\pi}{\mathrm{E}(|U|)+\mathrm{E}(|\xi|)}$ and $\lambda_T < \frac{\epsilon}{4\left\|f'_\xi\right\|_\infty \mathrm{E}(|U|)}$, this means*
$T > \max\left\{\left(\frac{4(\mathrm{E}(|U|)+\mathrm{E}(|\xi|))}{2\pi\epsilon}\right)^{\frac{1}{3c_2\cdot c}}, \exp\left(\frac{4c\left\|f'_\xi\right\|_\infty \mathrm{E}(|U|)}{\epsilon}\right)\right\}$, *then*

$$\mathrm{P}\left(\left|\hat{f}_\xi^T(x) - f_\xi(x)\right| > \epsilon\right) \leq \frac{8(2+c_T^2)}{T\lambda_T d_T c_T^4} + \frac{8}{T\lambda_T^3 d_T^3 c_T^2 \pi^2 \epsilon^2}\left(\frac{10-2\rho_\xi}{1-\rho_\xi}\right).$$

**Proof:**

$$\begin{aligned}
\left|\hat{f}_\xi^T(x) - f_\xi(x)\right| &\leq \left|\hat{f}_\xi^T(x) - R_T(x)\right| + |R_T(x) - I_T(x)| + |I_T(x) - f_\xi(x)| \\
&\leq \left|\hat{f}_\xi^T(x) - R_T(x)\right| + \frac{\epsilon}{2},
\end{aligned}$$

*where the last inequality is due to Lemma A.5 b) and c) and the fact, that $T$ is chosen large enough. Thus*

$$\mathrm{P}\left(\left|\hat{f}_\xi^T(x) - f_\xi(x)\right| > \epsilon\right) \leq \mathrm{P}\left(\left|\hat{f}_\xi^T(x) - R_T(x)\right| \geq \epsilon/2\right).$$

*The statement is now an immediate consequence of Lemma A.5 a).* □

**Proof of Theorem 2.2:**
Fix an arbitrary $\epsilon > 0$ and choose $T$ large enough, such that the condition of Lemma A.7 is fulfilled.

$$\begin{aligned}
&\mathrm{P}\left(\left|\hat{f}_\xi^T(x) - f_\xi(x)\right| > \epsilon\right) \\
&\quad\leq \frac{16}{Tc^2(\log T)^{-2}T^{-3c_2\cdot c}T^{-4c_2\cdot c}} + \frac{8T^{-2c_2\cdot c}}{Tc^2(\log T)^{-2}T^{-3c_2\cdot c}T^{-4c_2\cdot c}} \\
&\qquad+\frac{8}{Tc^6(\log T)^{-6}T^{-9c_2\cdot c}T^{-2c_2\cdot c}\pi^2\epsilon^2}\left(\frac{10-2\rho_\xi}{1-\rho_\xi}\right) \\
&\quad= \frac{16(\log T)^2}{c^2T^{1-7c_2\cdot c}} + \frac{8(\log T)^2}{c^2T^{1-5c_2\cdot c}} + \frac{8(\log T)^6}{c^6T^{1-11c_2\cdot c}\pi^2\epsilon^2}\left(\frac{10-2\rho_\xi}{1-\rho_\xi}\right) \\
&\quad\xrightarrow{T\to\infty} 0,
\end{aligned}$$

since $1 - 11c_2 \cdot c > 0$. □

## A.2 Proof of Proposition 2.3

To prove Proposition 2.3, we write the term $\frac{1}{T_2 M\lambda_T}\sum_{j=1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)X_{j+1,m}$ in the following way:

$$\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)X_{j+1,m}$$

$$= \frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(\xi_{j+1}+\epsilon_{j+1,m})$$

$$= \frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(m(\xi_j)+\eta_{j+1}+\epsilon_{j+1,m})$$

$$\begin{aligned}
&= \frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(m(\xi_j)-m(x)) \\
&\quad+\frac{m(x)}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right) \\
&\quad+\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\eta_{j+1} \\
&\quad+\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\epsilon_{j+1,m}
\end{aligned} \tag{A.4}$$

and study the four summands of (A.4) separately.
The estimator $\hat{\phi}_X^{T,2}$ is defined just like the estimator $\hat{\phi}_X^T$, but uses only the second part of the observation period, namely the variables $X_{j,m}$ with $j > T_1$.
Simply changing $T$ to $T_1$ or $T_2$ in the Lemmas A.1, A.2, A.3 and A.4, which were proved above, leads to the following estimations:

**A.8 Lemma.**
*If A1 and A2 are fulfilled, then*

*a)*

$$\mathrm{P}\left(\left|\hat{\phi}_X^{T,1}(\tau)-\phi_X(\tau)\right|>\epsilon\right)\le\frac{1}{T_1\epsilon^2}\left(\frac{3-\rho_\xi}{9-\rho_\xi}\right),$$

*b)*

$$\mathrm{P}\left(\left|\hat{\phi}_Z^{T,1}(\tau)-\phi_Z(\tau)\right|>\epsilon\right)\le\frac{1}{T_1\epsilon^2},$$

*c)*

$$\mathrm{P}\left(\left|\hat{\phi}_X^{T,2}(\tau) - \phi_X(\tau)\right| > \epsilon\right) \leq \frac{1}{T_2\epsilon^2}\left(\frac{3-\rho_\xi}{9-\rho_\xi}\right).$$

*If additional A5 is fulfilled and* $\tau \in [-\frac{1}{\lambda_T}, \frac{1}{\lambda_T}]$, *then*

*d)*

$$\mathrm{P}\left(\left|\tilde{\phi}_Z^{T,1}(\tau) - \phi_Z(\tau)\right| > \epsilon\right) \leq \frac{1}{T_1\epsilon^2} + \frac{4}{T_1 c_T^4},$$

*e)*

$$\mathrm{P}\left(\left|\hat{\phi}_\epsilon^{T,1}(\tau) - \phi_\epsilon(\tau)\right| > \epsilon\right) \leq \frac{1}{T_1 c_T^2\epsilon^2} + \frac{4}{T_1 c_T^4}.$$

**A.9 Lemma.**
*If A1, A2, A5, A6 and A7 are fulfilled and if all tuning parameters are chosen as in* (2.17), *then*

$$\frac{m(x)}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right) \xrightarrow[T\to\infty]{p} m(x)f_\xi(x).$$

**Proof:**
*Easy calculation shows*

$$\frac{m(x)}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right) = m(x)\cdot\hat{f}_\xi^{T,2}(x),$$

*with*

$$\hat{f}_\xi^{T,2}(x) := \frac{1}{2\pi}\sum_{j=1}^{S_T} e^{-i\tau_j x}\phi_U\left(\lambda_T\tau_j^T\right)\frac{\hat{\phi}_X^{T,2}\left(\tau_j^T\right)}{\hat{\phi}_\epsilon^{T,1}\left(\tau_j^T\right)}d_T.$$

*Like in Theorem 2.2, recall the preparing Lemmas A.5 - A.7, we can show, that if* $T$ *is large enough*

$$\mathrm{P}\left(\left|\hat{f}_\xi^{T,2}(x) - f_\xi(x)\right| > \epsilon\right) \leq \frac{8(2+c_T^2)}{T_1\lambda_T d_T c_T^4} + \frac{8}{T_2\lambda_T^3 d_T^3 c_T^2\pi^2\epsilon^2}\left(\frac{9-2\rho_\xi}{1-\rho_\xi}\right) + \frac{8}{T_1\lambda_T^3 d_T^3 c_T^2\pi^2\epsilon^2},$$

*we just have to take in mind, that* $\hat{\phi}_X^{T,2}$ *is based on the second part of the observation period and* $\hat{\phi}_\epsilon^{T,1}$ *is based on the first part of the observation period, and use the estimations in Lemma A.8.*
*The stochastic convergence is now an immediate consequence. Since* $m(x)$ *is just a constant, the statement is now clear.*

$\square$

Since we know now, that the second summand of (A.4) converges to $m(x)f(x)$ in probability, we have to show, that the other addends converge to zero in probability, which will complete the proof of Proposition 2.3. We will start with showing the stochastic convergence of the third and the fourth summand.

**A.10 Lemma.**
*If A1, A2, A5 and A7 are fulfilled, then*

*a)*

$$\mathrm{P}\left(\left|\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\epsilon_{j+1,m}\right|>\epsilon\right)\leq\frac{1}{T_2\lambda_T^3 d_T c_T^2}\cdot\frac{\sigma_\epsilon^2}{2\pi^2\epsilon^2}.$$

*b) If additionally all tuning parameters are chosen as in* (2.17) *and in* (2.21)*, then*

$$\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\epsilon_{j+1,m}\xrightarrow[T\to\infty]{p}0.$$

**Proof:**

*a)*

$$\begin{aligned}
&\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\epsilon_{j+1,m}\\
&\quad=\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\frac{\delta_T}{2\pi}\sum_{k=1}^{S_T}e^{-it_k^T\frac{x-X_{j,m}}{\lambda_T}}\frac{\phi_U(\tau_k^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_k^T)}\epsilon_{j+1,m}\\
&\quad=\frac{d_T}{2\pi T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\sum_{k=1}^{S_T}e^{-i\tau_k^T(x-X_{j,m})}\frac{\phi_U(\tau_k^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_k^T)}\epsilon_{j+1,m}.
\end{aligned}$$

*Thus*

$$\begin{aligned}
&\mathrm{P}\left(\left|\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\epsilon_{j+1,m}\right|>\epsilon\right)\\
&=\ \mathrm{P}\left(\left|\frac{d_T}{2\pi T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\sum_{k=1}^{S_T}e^{-i\tau_k^T(x-X_{j,m})}\frac{\phi_U(\tau_k^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_k^T)}\epsilon_{j+1,m}\right|>\epsilon\right)\\
&\leq\ \mathrm{P}\left(\max_{k\in\{1,\dots,S_T\}}\left|\frac{d_T}{2\pi T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}e^{-i\tau_k^T(x-X_{j,m})}\frac{\phi_U(\tau_k^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_k^T)}\epsilon_{j+1,m}\right|>\frac{\epsilon}{S_T}\right)\\
&=\ \mathrm{P}\left(\bigcup_{j=1}^{S_T}\left\{\left|\frac{d_T}{2\pi T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}e^{-i\tau_k^T(x-X_{j,m})}\frac{\phi_U(\tau_k^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_k^T)}\epsilon_{j+1,m}\right|>\frac{\epsilon}{S_T}\right\}\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{k=1}^{S_T} \mathrm{P}\left(\left|\frac{d_T}{2\pi T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M} e^{-i\tau_k^T(x-X_{j,m})}\frac{\phi_U(\tau_k^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_k^T)}\epsilon_{j+1,m}\right| > \frac{\epsilon}{S_T}\right)\\
&= S_T\cdot \mathrm{P}\left(\left|\sum_{j=T_1+1}^{T}\sum_{m=1}^{M} e^{-i\tau_1^T(x-X_{j,m})}\frac{\phi_U(\tau_1^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_1^T)}\epsilon_{j+1,m}\right| > \frac{\epsilon\cdot 2\pi\cdot T_2\cdot M}{S_T\cdot d_T}\right)\\
&\leq S_T\cdot \mathrm{Var}\left(\sum_{j=T_1+1}^{T}\sum_{m=1}^{M} e^{-i\tau_1^T(x-X_{j,m})}\frac{\phi_U(\tau_1^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_1^T)}\epsilon_{j+1,m}\right)\frac{S_T^2\cdot d_T^2}{\epsilon^2\cdot 4\pi^2\cdot T_2^2\cdot M^2}.
\end{aligned}
$$

*The last inequality is due to Chebyshev's inequality, which can be used since*

$$
\mathrm{E}\left(\sum_{j=T_1+1}^{T}\sum_{m=1}^{M} e^{-i\tau_1^T(x-X_{j,m})}\frac{\phi_U(\tau_1^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_1^T)}\epsilon_{j+1,m}\right) = 0, \tag{A.5}
$$

*which is due to the independence of* $\epsilon_{j+1,m}$ *of the other two factors. Because of* (A.5) *we get:*

$$
\begin{aligned}
&\mathrm{Var}\left(\sum_{j=T_1+1}^{T}\sum_{m=1}^{M} e^{-i\tau_1^T(x-X_{j,m})}\frac{\phi_U(\tau_1^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_1^T)}\epsilon_{j+1,m}\right)\\
&= \mathrm{E}\left(\left(\sum_{j=T_1+1}^{T}\sum_{m=1}^{M} e^{-i\tau_1^T(x-X_{j,m})}\frac{\phi_U(\tau_1^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_1^T)}\epsilon_{j+1,m}\right)^2\right)\\
&= \sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\mathrm{E}\left(\frac{\phi_U^2(\tau_1^T\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau_1^T)}\epsilon_{j+1,m}^2\right)\\
&\quad +2\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\sum_{n=m+1}^{M}\mathrm{E}\left(e^{-i\tau_1^T(\epsilon_{j,n}-\epsilon_{j,m})}\frac{\phi_U^2(\tau_1^T\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau_1^T)}\epsilon_{j+1,m}\epsilon_{j+1,n}\right)\\
&\quad +2\sum_{j=T_1+1}^{T}\sum_{l=j+1}^{T}\sum_{m,n=1}^{M}\mathrm{E}\left(e^{-i\tau_1^T(X_{l,n}-X_{j,m})}\frac{\phi_U^2(\tau_1^T\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau_1^T)}\epsilon_{j+1,m}\epsilon_{l+1,n}\right).
\end{aligned}
$$

*We now look separately at the three summands above. Let* $\tau\in[-\lambda_T^{-1},\lambda_T^{-1}]$*:*

$$
0\leq \mathrm{E}\left(\frac{\phi_U^2(\tau\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau)}\epsilon_{j+1,m}^2\right)\leq \mathrm{E}\left(\frac{2}{c_T^2}\epsilon_{j+1,m}^2\right) = \frac{2\sigma_\epsilon^2}{c_T^2}.
$$

*Let* $n>m$, $\tau\in[-\lambda_T^{-1},\lambda_T^{-1}]$*:*

$$
\begin{aligned}
\left|\mathrm{E}\left(e^{-i\tau_k^T(\epsilon_{j,n}-\epsilon_{j,m})}\frac{\phi_U^2(\tau\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau)}\epsilon_{j+1,m}\epsilon_{j+1,n}\right)\right| &\leq \mathrm{E}\left(\frac{2}{c_T^2}\left|\epsilon_{j+1,m}\epsilon_{j+1,n}\right|\right)\\
&= \frac{2}{c_T^2}\mathrm{E}\left(\left|\epsilon_{j+1,m}\right|\right)^2\leq\frac{2\sigma_\epsilon^2}{c_T^2}.
\end{aligned}
$$

*Let $l > j$, $\tau \in I\!R$:*

$$\begin{aligned}
&\mathrm{E}\left(e^{-i\tau(X_{l,n}-X_{j,m})}\frac{\phi_U^2(\tau\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau)}\epsilon_{j+1,m}\epsilon_{l+1,n}\right)\\
&\quad = \mathrm{E}\left(e^{-i\tau(2x-X_{j,m}-X_{l,n})}\frac{\phi_U^2(\tau\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau)}\epsilon_{j+1,m}\right)\mathrm{E}\left(\epsilon_{l+1,n}\right) = 0.
\end{aligned}$$

*So we get for $\tau \in [-\lambda_T^{-1}, \lambda_T^{-1}]$*

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{j=T_1+1}^{T}\sum_{m=1}^{M} e^{-i\tau_k^T(x-X_{j,m})}\frac{\phi_U(\tau\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau)}\epsilon_{j+1,m}\right)\\
&\quad \leq T_2M\frac{2\sigma_\epsilon^2}{c_T^2} + T_2M\frac{2\sigma_\epsilon^2}{c_T^2}T_2M(M-1)\frac{2\sigma_\epsilon^2}{c_T^2} = T_2M^2\frac{\sigma_\epsilon^2}{c_T^2},
\end{aligned}$$

*which leads to*

$$\begin{aligned}
&\mathrm{P}\left(\left|\frac{1}{T_2M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\epsilon_{j+1,m}\right| > \epsilon\right)\\
&\leq S_TT_2M^2\frac{2\sigma_\epsilon^2}{c_T^2}\cdot\frac{S_T^2\cdot d_T^2}{2\pi^2\epsilon^2T_2^2M^2} = \frac{S_T^3d_T^2}{T_2c_T^2}\cdot\frac{\sigma_\epsilon^2}{2\pi^2\epsilon^2} = \frac{d_T^2}{T_2\lambda_T^3d_T^3c_T^2}\cdot\frac{\sigma_\epsilon^2}{2\pi^2\epsilon^2}\\
&= \frac{1}{T_2\lambda_T^3d_Tc_T^2}\cdot\frac{\sigma_\epsilon^2}{4\pi^2\epsilon^2}.
\end{aligned}$$

b) *Fix an arbitrary $\epsilon > 0$, then it follows directly from a):*

$$\begin{aligned}
&\mathrm{P}\left(\left|\frac{1}{T_2M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\epsilon_{j+1,m}\right| > \epsilon\right)\\
&\leq \frac{1}{T_2\lambda_T^3d_Tc_T^2}\cdot\frac{\sigma_\epsilon^2}{2\pi^2\epsilon^2} = \frac{1}{T_2a_T\lambda_T^4c_T^2}\cdot\frac{\sigma_\epsilon^2}{2\pi^2\epsilon^2}\\
&\leq \frac{(\log T)^4}{(1-a)TT^{-3c_2\cdot c}cT^{-2c_2\cdot c}}\cdot\frac{\sigma_\epsilon^2}{2\pi^2\epsilon^2} = \frac{(\log T)^4}{T^{1-5c_2\cdot c}}\cdot\frac{c_1^2\sigma_\epsilon^2}{2(1-a)c\pi^2\epsilon^2}\\
&\stackrel{T\to\infty}{\longrightarrow} 0.
\end{aligned}$$

$\square$

**A.11 Lemma.**
*If A1, A2, A3, A5 and A7 are fulfilled, then*

*a)*

$$\mathrm{P}\left(\left|\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\eta_{j+1}\right|>\epsilon\right)\leq\frac{1}{T_2\lambda_T^3 d_T c_T^2}\cdot\frac{\sigma_\eta^2}{2\pi^2\epsilon^2}.$$

*b) If additionally all tuning parameters are chosen as in* (2.17) *and in* (2.21)*, then*

$$\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\eta_{j+1}\ \xrightarrow[T\to\infty]{p}\ 0.$$

**Proof:**

*a) The proof is almost the same as the one of Lemma A.10a). Just like there it can be shown, that*

$$\begin{aligned}&\mathrm{P}\left(\left|\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\eta_{j+1}\right|>\epsilon\right)\\ &\leq\ S_T\mathrm{Var}\left(\left|\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}e^{-i\tau_k^T(x-X_{j,m})}\frac{\phi_U(\tau_1^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_1^T)}\eta_{j+1}\right|\right)\frac{S_T^2\cdot d_T^2}{\epsilon^2\cdot 4\pi^2\cdot T_2^2\cdot M^2}\end{aligned}$$

*and*

$$\begin{aligned}&\mathrm{Var}\left(\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}e^{-i\tau_1^T(x-X_{j,m})}\frac{\phi_U(\tau_1^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_1^T)}\eta_{j+1}\right)\\
&=\ \mathrm{E}\left(\left(\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}e^{-i\tau_1^T(x-X_{j,m})}\frac{\phi_U(\tau_1^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_1^T)}\eta_{j+1}\right)\right.\\
&\qquad\left.\left(\sum_{l=T_1+1}^{T}\sum_{n=1}^{M}e^{+i\tau_1^T(x-X_{l,n})}\frac{\phi_U(\tau_1^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_1^T)}\eta_{l+1}\right)\right)\\
&=\ \sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\mathrm{E}\left(\frac{\phi_U^2(\tau_1^T\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau_1^T)}\eta_{j+1}^2\right)\\
&\quad+2\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\sum_{n=m+1}^{M}\mathrm{E}\left(e^{-i\tau_1^T(\epsilon_{j,n}-\epsilon_{j,m})}\frac{\phi_U^2(\tau_1^T\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau_1^T)}\eta_{j+1}^2\right)\\
&\quad+2\sum_{j=T_1+1}^{T}\sum_{l=j+1}^{T}\sum_{m,n=1}^{M}\mathrm{E}\left(e^{-i\tau_1^T(X_{l,n}-X_{j,m})}\frac{\phi_U^2(\tau_1^T\lambda_T)}{\tilde{\phi}_Z^{T,1}(\tau_1^T)}\eta_{j+1}\eta_{l+1}\right)\\
&\leq\ T_2M\frac{2\sigma_\eta^2}{c_T^2}+T_2M(M-1)\frac{2\sigma_\eta^2}{c_T^2}=T_2M^2\frac{2\sigma_\eta^2}{c_T^2},\end{aligned}$$

*where the inequality can be shown just as in the proof of Lemma A.10a) and so we achieve:*

$$\mathrm{P}\left(\left|\frac{1}{T_2 \cdot M \cdot \lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\epsilon_{j+1,m}\right| > \epsilon\right) \leq \frac{1}{T_2\lambda_T^3 d_T c_T^2}\cdot\frac{\sigma_\eta^2}{2\pi^2\epsilon^2}.$$

b) *Since the time variables on the right hand side of the inequality in a) are the same as the ones on the right hand side of Lemma A.10a), b) can be shown just like Lemma A.10b).*

□

To prove the stochastic convergence of the first summand of (A.4), we first have to introduce the auxiliary random variables $W_T(x)$ and $J_T(x)$ by

$$W_T(x) := \frac{d_T}{T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\sum_{k=1}^{S_T}\frac{1}{2\pi}e^{-i\tau_j^T(x-X_{j,m})}\frac{\phi_U(\tau_k^T\lambda_T)}{\phi_\epsilon(\tau_k^T)}\left(m(\xi_j)-m(x)\right),$$

$$J_T(x) := \frac{1}{T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\int_{\mathbb{R}}\frac{1}{2\pi}e^{-i\tau(x-X_{j,m})}\frac{\phi_U(\tau\lambda_T)}{\phi_\epsilon(\tau)}\,d\tau\left(m(\xi_j)-m(x)\right),$$

divide the first summand into three parts

$$\begin{aligned}
&\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\left(m(\xi_j)-m(x)\right)\\
&\quad=\left(\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\left(m(\xi_j)-m(x)\right)-W_T(x)\right)\\
&\qquad+\left(W_T(x)-J_T(x)\right)+J_T(x)
\end{aligned}$$

and show that all these three summands converge to zero in probability, which immediately implies the following Lemma:

**A.12 Lemma.**
*If A1, A2, A4, A5, A6 and A7 are fulfilled and if all tuning parameters are chosen as in* (2.17) *and in* (2.21), *then*

$$\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)\left(m(\xi_j)-m(x)\right)\xrightarrow[T\to\infty]{p}0.$$

**A.13 Lemma.**
*If A1, A2 and A5 are fulfilled, then*

*a)*

$$
\begin{aligned}
&\mathrm{P}\left(\left|\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(m(\xi_j)-m(x))-W_T(x)\right|>\epsilon\right)\\
&\quad\leq \frac{1}{\pi\epsilon\lambda_T c_T^3\sqrt{T_1}}\mathrm{E}\left(|m(\xi_T)-m(x)|\right).
\end{aligned}
$$

*b) If additionally all tuning parameters are chosen as in* (2.17) *and in* (2.21), *then*

$$
\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(m(\xi_j)-m(x))-W_T(x)\xrightarrow[T\to\infty]{p}0.
$$

**Proof:**

*a)*

$$
\begin{aligned}
&\mathrm{E}\left(\left|\frac{1}{T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(m(\xi_j)-m(x))-W_T(x)\right|\right)\\
&= \mathrm{E}\left(\left|\frac{d_T}{T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\sum_{k=1}^{S_T}\frac{1}{2\pi}e^{-i\tau_k^T(x-X_{j,m})}\frac{\phi_U(\tau_k^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_k^T)}(m(\xi_j)-m(x))-W_T(x)\right|\right)\\
&= \mathrm{E}\left(\left|\frac{d_T}{T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\sum_{k=1}^{S_T}\frac{1}{2\pi}e^{-i\tau_k^T(x-X_{j,m})}\left[\frac{\phi_U(\tau_k^T\lambda_T)}{\hat{\phi}_\epsilon^{T,1}(\tau_k^T)}-\frac{\phi_U(\tau_k^T\lambda_T)}{\phi_\epsilon(\tau_k^T)}\right](m(\xi_j)-m(x))\right|\right)\\
&\leq \frac{d_T}{2\pi}\sum_{k=1}^{S_T}\mathrm{E}\left(\left|\frac{1}{\hat{\phi}_\epsilon^{T,1}(\tau_k^T)}-\frac{1}{\phi_\epsilon(\tau_k^T)}\right|\right)\mathrm{E}\left(\left|e^{-i\tau_k^T(x-X_{T,M})}\right|\left|\phi_U(\tau_k^T\lambda_T)\right|\left|m(\xi_T)-m(x)\right|\right)
\end{aligned}
$$

*Let* $\tau\in[-\lambda_T^{-1},\lambda_T^{-1}]$. *Recall, that* $\phi_\epsilon(\tau)>c_T$ *on* $[-\lambda_T^{-1},\lambda_T^{-1}]$ *and* $\hat{\phi}_\epsilon^{T,1}(\tau)>\frac{c_T}{2}$ $\forall\,\tau\in I\!R$*:*

$$
\begin{aligned}
&\mathrm{E}\left(\left|\frac{1}{\hat{\phi}_\epsilon^{T,1}(\tau)}-\frac{1}{\phi_\epsilon(\tau)}\right|\right)\\
&= \mathrm{E}\left(\left|\frac{\phi_\epsilon(\tau)-\hat{\phi}_\epsilon^{T,1}(\tau)}{\hat{\phi}_\epsilon^{T,1}(\tau)\phi_\epsilon(\tau)}\right|\right)\leq\mathrm{E}\left(\left|\frac{\phi_\epsilon(\tau)+\hat{\phi}_\epsilon^{T,1}(\tau)}{\frac{3c_T}{2}}\right|\cdot\left|\frac{\phi_\epsilon(\tau)-\hat{\phi}_\epsilon^{T,1}(\tau)}{\hat{\phi}_\epsilon^{T,1}(\tau)\phi_\epsilon(\tau)}\right|\right)\\
&= \frac{2}{3c_T}\mathrm{E}\left(\frac{\left|\phi_Z(\tau)-\tilde{\phi}_Z^{T,1}(\tau)\right|}{\hat{\phi}_\epsilon^{T,1}(\tau)\phi_\epsilon(\tau)}\right)\leq\frac{2}{3c_T}\mathrm{E}\left(\frac{\left|\phi_Z(\tau)-\tilde{\phi}_Z^{T,1}(\tau)\right|}{c_T^2/\sqrt{2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{\sqrt{8}}{3c_T^3}\sqrt{\mathrm{E}\left(\left|\phi_Z(\tau)-\tilde{\phi}_Z^{T,1}(\tau)\right|^2\right)} \leq \frac{\sqrt{8}}{3c_T^3}\sqrt{\mathrm{E}\left(\left(\mathcal{R}e\left(\phi_Z(\tau)\right)-\mathcal{R}e\left(\hat{\phi}_Z^{T,1}(\tau)\right)\right)^2\right)}\\
&\leq \frac{\sqrt{8}}{3c_T^3}\sqrt{\mathrm{E}\left(\left|\phi_Z(\tau)-\hat{\phi}_Z^{T,1}(\tau)\right|^2\right)} = \frac{\sqrt{8}}{3c_T^3}\sqrt{\mathrm{Var}\left(\hat{\phi}_Z^{T,1}(\tau)\right)} = \frac{\sqrt{8}}{3c_T^3}\cdot\frac{1}{\sqrt{T_1}}
\end{aligned}
$$

*Thus*

$$
\begin{aligned}
&\mathrm{E}\left(\left|\frac{1}{T_2M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(m(\xi_j)-m(x))-W_T(x)\right|\right)\\
&\leq \frac{d_T}{2\pi}\sum_{k=1}^{S_T}\frac{\sqrt{8}}{3c_T^3}\cdot\frac{1}{\sqrt{T_1}}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\\
&= \frac{d_T}{2\pi}2(\lambda_T d_T)^{-1}\frac{\sqrt{8}}{3c_T^3}\cdot\frac{1}{\sqrt{T_1}}\mathrm{E}\left(|m(\xi)-m(x)|\right)\\
&\leq \frac{1}{\pi\lambda_T c_T^3\sqrt{T_1}}\mathrm{E}\left(|m(\xi)-m(x)|\right)
\end{aligned}
$$

*and by using Markov's inequality we get:*

$$
\begin{aligned}
&\mathrm{P}\left(\left|\frac{1}{T_2M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(m(\xi_j)-m(x))-W_T(x)\right|>\epsilon\right)\\
&\leq \frac{1}{\epsilon}\mathrm{E}\left(\left|\frac{1}{T_2M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(m(\xi_j)-m(x))-W_T(x)\right|\right)\\
&\leq \frac{1}{\pi\epsilon\lambda_T c_T^3\sqrt{T_1}}\mathrm{E}\left(|m(\xi_T)-m(x)|\right).
\end{aligned}
$$

*b) Fix an arbitrary* $\epsilon>0$ *and let* $T\geq\frac{2}{a}$*. From Lemma A.13 we get*

$$
\begin{aligned}
&\mathrm{P}\left(\left|\frac{1}{T_2M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\hat{K}_T^1\left(\frac{x-X_{j,m}}{\lambda_T}\right)(m(\xi_j)-m(x))-W_T(x)\right|>\epsilon\right)\\
&\leq \frac{1}{\pi\epsilon\lambda_T c_T^3\sqrt{T_1}}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\\
&\leq \frac{c^3}{\pi\epsilon c(\log T)^{-1}T^{-3c_2\cdot c}\sqrt{\frac{aT}{2}}}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\\
&= \frac{c^3\sqrt{2}\log T}{\pi\epsilon\sqrt{a}cT^{\frac{1}{2}-3c_2\cdot c}}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\\
&\overset{T\to\infty}{\longrightarrow} 0.
\end{aligned}
$$

$\square$

**A.14 Lemma.**
*If A1, A2 and A5 are fulfilled, then*

*a)*

$$\begin{aligned}
&\mathrm{P}\left(|W_T(x) - J_T(x)| > \epsilon\right)\\
\leq\ & \frac{d_T}{2\pi\epsilon\lambda_T c_T}\mathrm{E}\left(|m(\xi_T) - m(x)|\,|x - X_{T,M}|\right) + \frac{d_T}{2\pi\epsilon c_T^2}\mathrm{E}\left(|m(\xi_T) - m(x)|\right)\mathrm{E}\left(|U|\right)\\
&+\frac{d_T}{2\pi\epsilon\lambda_T c_T^2}\mathrm{E}\left(|m(\xi_T) - m(x)|\right)\mathrm{E}\left(|\epsilon_{T,M}|\right).
\end{aligned}$$

*b) If additionally all tuning parameters are chosen as in* (2.17) *and in* (2.21), *then*

$$W_T(x) - J_T(x) \xrightarrow[T\to\infty]{p} 0.$$

**Proof:**

*a)*

$$\begin{aligned}
&|W_T(x) - J_T(x)|\\
=\ & \Bigg|\frac{d_T}{T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\sum_{k=1}^{S_T}\frac{1}{2\pi}e^{-i\tau_k^T(x - X_{j,m})}\frac{\phi_U(\tau_k^T\lambda_T)}{\phi_\epsilon(\tau_k^T)}\left(m(\xi_j) - m(x)\right)\\
&-\frac{1}{T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\int_{\mathbb{R}}\frac{1}{2\pi}e^{-i\tau(x - X_{j,m})}\frac{\phi_U(\tau\lambda_T)}{\phi_\epsilon(\tau)}\left(m(\xi_j) - m(x)\right)\,d\tau\Bigg|\\
\leq\ & \frac{1}{T_2 M}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\frac{1}{2}d_T^2 S_T\sup_{\tau\in\mathbb{R}}\left|\frac{\partial}{\partial\tau}\left(\frac{1}{2\pi}e^{-i\tau(x - X_{j,m})}\frac{\phi_U(\tau\lambda_T)}{\phi_\epsilon(\tau)}\left(m(\xi_j) - m(x)\right)\right)\right|\\
=\ & \frac{d_T}{2\pi T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}|m(\xi_j) - m(x)|\\
&\cdot\sup_{\tau\in\mathbb{R}}\Bigg|-i(x - X_{j,m})e^{-i\tau(x - X_{j,m})}\frac{\phi_U(\tau\lambda_T)}{\phi_\epsilon(\tau)}\\
&\qquad+e^{-i\tau(x - X_{j,m})}\frac{\phi_U'(\tau\lambda_T)\lambda_T\phi_\epsilon(\tau) - \phi_U(\tau\lambda_T)\phi_\epsilon'(\tau)}{\phi_\epsilon^2(\tau)}\Bigg|\\
\leq\ & \frac{d_T}{2\pi T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}|m(\xi_j) - m(x)|\left(|x - X_{j,m}|\frac{1}{c_T} + \frac{\mathrm{E}\left(|U|\right)\lambda_T + \mathrm{E}\left(|\epsilon_{T,M}|\right)}{c_T^2}\right)
\end{aligned}$$

*Thus*

$$\mathrm{P}\left(|W_T(x) - J_T(x)| > \epsilon\right) \leq \frac{1}{\epsilon}\mathrm{E}\left(|W_T(x) - J_T(x)|\right)$$

$$
\begin{aligned}
&\leq \frac{1}{\epsilon}\mathrm{E}\left(\frac{d_T}{2\pi T_2 M\lambda_T}\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}|m(\xi_j)-m(x)|\right.\\
&\qquad\qquad\left.\cdot\left(|x-X_{j,m}|\frac{1}{c_T}+\frac{\mathrm{E}\left(|U|\right)\lambda_T+\mathrm{E}\left(|\epsilon_{T,M}|\right)}{c_T^2}\right)\right)\\
&\leq \frac{d_T}{2\pi\epsilon\lambda_T}\mathrm{E}\left(|m(\xi_T)-m(x)|\left(|x-X_{T,M}|\frac{1}{c_T}+\frac{\mathrm{E}\left(|U|\right)\lambda_T+\mathrm{E}\left(|\epsilon_{T,M}|\right)}{c_T^2}\right)\right)\\
&= \frac{d_T}{2\pi\epsilon\lambda_T c_T}\mathrm{E}\left(|m(\xi_T)-m(x)|\,|x-X_{T,M}|\right)+\frac{d_T}{2\pi\epsilon c_T^2}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\mathrm{E}\left(|U|\right)\\
&\quad+\frac{d_T}{2\pi\epsilon\lambda_T c_T^2}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\mathrm{E}\left(|\epsilon_{T,M}|\right).
\end{aligned}
$$

*b)*

$$
\begin{aligned}
&\mathrm{P}\left(|W_T(x)-J_T(x)|>\epsilon\right)\\
&\quad\leq \frac{d_T}{2\pi\epsilon\lambda_T c_T}\mathrm{E}\left(|m(\xi_T)-m(x)|\,|x-X_{T,M}|\right)\\
&\qquad+\frac{d_T}{2\pi\epsilon c_T^2}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\mathrm{E}\left(|U|\right)\\
&\qquad+\frac{d_T}{2\pi\epsilon\lambda_T c_T^2}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\mathrm{E}\left(|\epsilon_{T,M}|\right)\\
&\quad= \frac{c_1 T^{-3c_2\cdot c}}{2\pi\epsilon T^{-c_2\cdot c}}\mathrm{E}\left(|m(\xi_T)-m(x)|\,|x-X_{T,M}|\right)\\
&\qquad+\frac{c_1^2 T^{-3c_2\cdot c}\log T}{2\pi\epsilon T^{-2c_2\cdot c}}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\mathrm{E}\left(|U|\right)\\
&\qquad+\frac{c_1^2 T^{-3c_2\cdot c}}{2\pi\epsilon T^{-2c_2\cdot c}}\mathrm{E}\left(|m(\xi_T)-m(x)|\right)\mathrm{E}\left(|\epsilon_{T,M}|\right)\\
&\quad\overset{T\to\infty}{\longrightarrow} 0.
\end{aligned}
$$

□

**A.15 Lemma.**
*If A1, A2, A4, A5, A6 and A7 are fulfilled, then*

*a)*

$$
\left|\mathrm{E}\left(J_T(x)\right)\right| \leq \lambda_T^2\left(m'(x)\left\|f_\xi'\right\|_\infty+\frac{1}{2}\left\|m''\right\|_\infty f_\xi(x)\right)+\lambda_T^3\frac{1}{2}\left\|m''\right\|_\infty\left\|f_\xi'\right\|_\infty\mathrm{E}\left(|U|^3\right),
$$

*b)*

$$
\begin{aligned}
\mathrm{Var}\left(J_T(x)\right) \leq \frac{1}{\pi^2 T_2\lambda_T^2}&\left(\left[\frac{1}{Mc_T^2}+\frac{M-1}{M}\right]\mathrm{E}\left((m(\xi_j)-m(x))^2\right)\right.\\
&\left.+8\frac{2+\delta}{\delta}\frac{1}{1-\rho_\xi^{\frac{\delta}{2+\delta}}}\left\|m(\xi_j)-m(x)\right\|_{2+\delta}^2\right).
\end{aligned}
$$

**Proof:**

*a)*

$$\begin{aligned}
\mathrm{E}\left(J_T(x)\right) &= \frac{1}{2\pi}\mathrm{E}\left(\int_{\mathbb{R}} e^{-i\tau(x-X_{T,M})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_T)-m(x))\right)\\
&= \frac{1}{2\pi\lambda_T}\mathrm{E}\left(\int_{\mathbb{R}} e^{-i\frac{\omega}{\lambda_T}(x-\xi_T-\epsilon_{T,M})}\frac{\phi_U(\omega)}{\phi_\epsilon\left(\omega/\lambda_T\right)}\,d\omega(m(\xi_T)-m(x))\right)\\
&= \frac{1}{2\pi\lambda_T}\int_{\mathbb{R}^3} e^{-i\frac{\omega}{\lambda_T}(x-u-v)}\frac{\phi_U(\omega)}{\phi_\epsilon\left(\omega/\lambda_T\right)}(m(u)-m(x))f_\xi(u)f_\epsilon(v)\,du\,dv\,d\omega\\
&= \frac{1}{2\pi\lambda_T}\int_{\mathbb{R}^2} e^{-i\frac{\omega}{\lambda_T}(x-u)}\phi_U(\omega)(m(u)-m(x))f_\xi(u)\,du\,d\omega\\
&= \frac{1}{\lambda_T}\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\frac{1}{2\pi}e^{-i\omega\frac{x-u}{\lambda_T}}\phi_U(\omega)\,d\omega\right)(m(u)-m(x))f_\xi(u)\,du\\
&= \frac{1}{\lambda_T}\int_{\mathbb{R}} K\left(\frac{x-u}{\lambda_T}\right)(m(u)-m(x))f_\xi(u)\,du\\
&= \int_{\mathbb{R}} K(v)(m(x-v\lambda_T)-m(x))f_\xi(x-v\lambda_T)\,dv\\
&= \int_{\mathbb{R}} K(v)\left(-v\lambda_T m'(x)+\frac{1}{2}v^2\lambda_T^2 m''(\hat{x}_1)\right)\left(f_\xi(x)-v\lambda_T f_\xi'(\hat{x}_2)\right)dv\\
&= -\int_{\mathbb{R}} K(v)v\lambda_T m'(x)f_\xi(x)\,dv+\int_{\mathbb{R}} K(v)v^2\lambda_T^2 m'(x)f_\xi'(\hat{x}_2)\,dv\\
&\quad+\int_{\mathbb{R}} K(v)\frac{1}{2}v^2\lambda_T^2 m''(\hat{x}_1)f_\xi(x)\,dv-\int_{\mathbb{R}} K(v)\frac{1}{2}v^3\lambda_T^3 m''(\hat{x}_1)f_\xi'(\hat{x}_2)\,dv,
\end{aligned}$$

*where $\hat{x}_1$ and $\hat{x}_2$ are interim values, which are dependent on $v$. Since $K$ is symmetric, the first summand is equal to zero and we get:*

$$\begin{aligned}
\left|\mathrm{E}\left(J_T(x)\right)\right| &\leq \lambda_T^2\left(m'(x)\left\|f_\xi'\right\|_\infty+\frac{1}{2}\left\|m''\right\|_\infty f_\xi(x)\right)\\
&\quad+\lambda_T^3\frac{1}{2}\left\|m''\right\|_\infty\left\|f_\xi'\right\|_\infty \mathrm{E}\left(|U|^3\right).
\end{aligned}$$

*b)*

$$\begin{aligned}
&\mathrm{Var}\left(J_T(x)\right)\\
&= \frac{1}{4\pi^2T_2^2M^2}\Bigg[\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\mathrm{Var}\left(\int_{\mathbb{R}} e^{-i\tau(x-X_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x))\right)\\
&\quad+2\sum_{j=T_1+1}^{T}\sum_{m=1}^{M}\sum_{n=m+1}^{M}\mathrm{Cov}\left(\int_{\mathbb{R}} e^{-i\tau(x-X_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x)),\right.
\end{aligned}$$

$$\int_{\mathbb{R}} e^{-i\tau(x-X_{j,n})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x))\Big)$$

$$+2\sum_{j=T_1+1}^{T}\sum_{k=j+1}^{T}\sum_{m,n=1}^{M}\mathrm{Cov}\Big(\int_{\mathbb{R}} e^{-i\tau(x-X_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x))$$

$$\int_{\mathbb{R}} e^{-i\tau(x-X_{k,n})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_k)-m(x))\Big)\Big].$$

*Let* $j\in\{T_1+1,\ldots,T\}$ *and* $m\in\{1,\ldots,M\}$*:*

$$\begin{aligned}&\mathrm{Var}\left(\int_{\mathbb{R}} e^{-i\tau(x-X_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x))\right)\\&\qquad\le\ \mathrm{E}\left(\left(\frac{2}{\lambda_T}\cdot\frac{1}{c_T}\right)^2(m(\xi_j)-m(x))^2\right)\end{aligned}$$

*Let* $j\in\{T_1+1,\ldots,T\}$ *and* $m,n\in\{1,\ldots,M\},m<n$*:*

$$\begin{aligned}&\mathrm{Cov}\Big(\int_{\mathbb{R}} e^{-i\tau(x-X_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x)),\\&\qquad\int_{\mathbb{R}} e^{-i\tau(x-X_{j,n})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x))\Big)\\
=\ &\mathrm{E}\left(\int_{\mathbb{R}^2} e^{-i\tau(x-X_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\frac{\phi_U(\lambda_T\omega)}{\phi_\epsilon(\omega)}(m(\xi_j)-m(x))^2e^{i\omega(x-X_{j,n})}\,d\tau\,d\omega\right)\\
&-\mathrm{E}\left(\int_{\mathbb{R}} e^{-i\tau(x-X_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x))\right)\\
&\quad\cdot\mathrm{E}\left(\int_{\mathbb{R}} e^{i\tau(x-X_{j,n})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x))\right)\\
=\ &\int_{\mathbb{R}^2}\mathrm{E}\left(e^{-i\tau(x-\xi_j-\epsilon_{j,m})}e^{i\omega(x-\xi_j-\epsilon_{j,n})}(m(\xi_j)-m(x))^2\right)\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\frac{\phi_U(\lambda_T\omega)}{\phi_\epsilon(\omega)}\,d\tau\,d\omega\\
&-\int_{\mathbb{R}}\mathrm{E}\left(e^{-i\tau(x-\xi_j-\epsilon_{j,m})}(m(\xi_j)-m(x))\right)\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau\\
&\quad\cdot\int_{\mathbb{R}}\mathrm{E}\left(e^{i\tau(x-\xi_j-\epsilon_{j,n})}(m(\xi_j)-m(x))\right)\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau\\
=\ &\int_{\mathbb{R}^2}\mathrm{E}\left(e^{i\tau\epsilon_{j,m}}\right)\mathrm{E}\left(e^{-i\omega\epsilon_{j,n}}\right)\mathrm{E}\left(e^{-i\tau(x-\xi_j)}e^{i\omega(x-\xi_j)}(m(\xi_j)-m(x))^2\right)\\
&\qquad\cdot\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\frac{\phi_U(\lambda_T\omega)}{\phi_\epsilon(\omega)}\,d\tau\,d\omega\\
&-\int_{\mathbb{R}}\mathrm{E}\left(e^{i\tau\epsilon_{j,m}}\right)\mathrm{E}\left(e^{-i\tau(x-\xi_j)}(m(\xi_j)-m(x))\right)\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau\\
&\quad\cdot\int_{\mathbb{R}}\mathrm{E}\left(e^{-i\tau\epsilon_{j,n}}\right)\mathrm{E}\left(e^{i\tau(x-\xi_j)}(m(\xi_j)-m(x))\right)\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau\end{aligned}$$

$$
\begin{aligned}
= & \int_{\mathbb{R}^2} \phi_\epsilon(-\tau)\phi_\epsilon(\omega)\mathrm{E}\left(e^{-i\tau(x-\xi_j)}e^{i\omega(x-\xi_j)}(m(\xi_j)-m(x))^2\right) \\
& \cdot \frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\frac{\phi_U(\lambda_T\omega)}{\phi_\epsilon(\omega)}\,d\tau\,d\omega \\
& - \int_{\mathbb{R}} \phi_\epsilon(-\tau)\mathrm{E}\left(e^{-i\tau(x-\xi_j)}(m(\xi_j)-m(x))\right)\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau \\
& \cdot \int_{\mathbb{R}} \phi_\epsilon(\tau)\mathrm{E}\left(e^{i\tau(x-\xi_j)}(m(\xi_j)-m(x))\right)\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau \\
= & \ \mathrm{E}\left(\int_{\mathbb{R}^2} e^{-i\tau(x-\xi_j)}e^{i\omega(x-\xi_j)}(m(\xi_j)-m(x))^2\phi_U(\lambda_T\tau)\phi_U(\lambda_T\omega)\,d\tau\,d\omega\right) \\
& -\mathrm{E}\left(\int_{\mathbb{R}} e^{-i\tau(x-\xi_j)}(m(\xi_j)-m(x))\phi_U(\lambda_T\tau)\,d\tau\right) \\
& \cdot\mathrm{E}\left(\int_{\mathbb{R}} e^{i\tau(x-\xi_j)}(m(\xi_j)-m(x))\phi_U(\lambda_T\tau)\,d\tau\right) \\
= & \ \mathrm{Var}\left(\int_{\mathbb{R}} e^{-i\tau(x-\xi_j)}\phi_U(\lambda_T\tau)\,d\tau(m(\xi_j)-m(x))\right) \\
\leq & \ \mathrm{E}\left((m(\xi_j)-m(x))^2\left[\frac{2}{\lambda_T}\right]^2\right)
\end{aligned}
$$

*Let* $j,k \in \{T_1+1,\ldots,T\}, j<k$ *and* $m,n \in \{1,\ldots,M\}$*:*

$$
\begin{aligned}
& \Bigg|\mathrm{Cov}\left(\int_{\mathbb{R}} e^{-i\tau(x-X_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_j)-m(x)),\right. \\
& \qquad \left.\int_{\mathbb{R}} e^{-i\tau(x-X_{k,n})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}\,d\tau(m(\xi_k)-m(x))\right)\Bigg| \\
= & \ \mathrm{E}\left(\int_{\mathbb{R}^2} e^{-i\tau(x-\xi_j-\epsilon_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_j)-m(x))\right. \\
& \qquad \left.\cdot e^{i\omega(x-\xi_k-\epsilon_{k,n})}\frac{\phi_U(\lambda_T\omega)}{\phi_\epsilon(\omega)}(m(\xi_k)-m(x))\,d\tau\,d\omega\right) \\
& -\mathrm{E}\left(\int_{\mathbb{R}} e^{-i\tau(x-\xi-\epsilon_{j,m})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_j)-m(x))\,d\tau\right) \\
& \cdot\mathrm{E}\left(\int_{\mathbb{R}} e^{i\tau(x-\xi_k-\epsilon_{k,n})}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_k)-m(x))\,d\tau\right) \\
= & \int_{\mathbb{R}^2} \mathrm{E}\left(e^{i\tau\epsilon_{j,m}}\right)\mathrm{E}\left(e^{-i\omega\epsilon_{k,n}}\right) \\
& \cdot\mathrm{E}\left(e^{-i\tau(x-\xi_j)}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_j)-m(x))e^{i\omega(x-\xi_k)}\frac{\phi_U(\lambda_T\omega)}{\phi_\epsilon(\omega)}(m(\xi_k)-m(x))\right) \\
& d\tau\,d\omega \\
& -\int_{\mathbb{R}} \mathrm{E}\left(e^{i\tau\epsilon_{j,m}}\right)\mathrm{E}\left(e^{-i\tau(x-\xi_j)}\frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_j)-m(x))\right)\,d\tau
\end{aligned}
$$

$$
\begin{aligned}
& \cdot \int_{\mathbb{R}} \mathrm{E}\left(e^{-i\tau\epsilon_{k,n}}\right) \mathrm{E}\left(e^{i\tau(x-\xi_k)} \frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_k)-m(x))\right) d\tau \\
= & \int_{\mathbb{R}^2} \phi_\epsilon(\omega)\phi_\epsilon(-\tau) \\
& \cdot \mathrm{E}\left(e^{-i\tau(x-\xi_j)} \frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_j)-m(x)) e^{i\omega(x-\xi_k)} \frac{\phi_U(\lambda_T\omega)}{\phi_\epsilon(\omega)}(m(\xi_k)-m(x))\right) d\tau\, d\omega \\
& - \int_{\mathbb{R}} \phi_\epsilon(\tau)\mathrm{E}\left(e^{-i\tau(x-\xi_j)} \frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_j)-m(x))\right) d\tau \\
& \cdot \int_{\mathbb{R}} \phi_\epsilon(-\tau)\mathrm{E}\left(e^{i\tau(x-\xi_k)} \frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_k)-m(x))\right) d\tau \\
= & \ \mathrm{E}\left(\int_{\mathbb{R}^2} \phi_\epsilon(\tau) e^{-i\tau(x-\xi_j)} \frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_j)-m(x)) \right. \\
& \qquad \left. \phi_\epsilon(\omega) e^{i\omega(x-\xi_k)} \frac{\phi_U(\lambda_T\omega)}{\phi_\epsilon(\omega)}(m(\xi_k)-m(x))\, d\tau\, d\omega\right) \\
& - \mathrm{E}\left(\int_{\mathbb{R}} \phi_\epsilon(\tau) e^{-i\tau(x-\xi_j)} \frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_j)-m(x))\, d\tau\right) \\
& \cdot \mathrm{E}\left(\int_{\mathbb{R}} \phi_\epsilon(\tau) e^{i\tau(x-\xi_k)} \frac{\phi_U(\lambda_T\tau)}{\phi_\epsilon(\tau)}(m(\xi_k)-m(x))\, d\tau\right) \\
= & \ \mathrm{Cov}\left(\int_{\mathbb{R}} e^{-i\tau(x-\xi_j)} \phi_U(\lambda_T\tau)(m(\xi_j)-m(x))\, d\tau, \right. \\
& \qquad \left. \int_{\mathbb{R}} e^{-i\tau(x-\xi_k)} \phi_U(\lambda_T\tau)(m(\xi_k)-m(x))\, d\tau\right) \\
\leq & \ 2\frac{2+\delta}{\delta}\Big(2\alpha(\sigma(\xi_j),\sigma(\xi_k))\Big)^{\frac{\delta}{2+\delta}} \left\|\int_{\mathbb{R}} e^{-i\tau(x-\xi_j)}\phi_U(\lambda_T\tau)\, d\tau(m(\xi_j)-m(x))\right\|_{2+\delta}^2 \\
\leq & \ 4\frac{2+\delta}{\delta}\Big(\alpha_\xi(k-j)\Big)^{\frac{\delta}{2+\delta}} \left(\frac{2}{\lambda_T}\right)^2 \|m(\xi_j)-m(x)\|_{2+\delta}^2 \\
\leq & \ 4\frac{2+\delta}{\delta}\left(\rho_\xi^{k-j}\right)^{\frac{\delta}{2+\delta}} \left(\frac{2}{\lambda_T}\right)^2 \|m(\xi_j)-m(x)\|_{2+\delta}^2 \\
= & \ 4\frac{2+\delta}{\delta}\left(\rho_\xi^{\frac{\delta}{2+\delta}}\right)^{k-j} \left(\frac{2}{\lambda_T}\right)^2 \|m(\xi_j)-m(x)\|_{2+\delta}^2 ,
\end{aligned}
$$

*where the last but two inequality is due to Corollary 1.1 of Bosq (1996) [21] with* $p = \frac{2+\delta}{\delta}$ *and* $q = r = 2+\delta$. *Thus*

$$
\begin{aligned}
& \mathrm{Var}\,(J_T(x)) \\
\leq & \ \frac{1}{4\pi^2 T_2^2 M^2}\left[T_2 M\, \mathrm{E}\left(\left(\frac{2}{\lambda_T}\cdot\frac{1}{c_T}\right)^2 (m(\xi_j)-m(x))^2\right)\right.
\end{aligned}
$$

$$
\begin{aligned}
&\qquad +T_2 M(M-1)\,\mathrm{E}\left(\left(\frac{2}{\lambda_T}\right)^2 (m(\xi_j)-m(x))^2\right)\\
&\qquad +2M^2 \sum_{j=T_1+1}^{T} \sum_{k=j+1}^{T} 4\frac{2+\delta}{\delta}\left(\rho_\xi^{\frac{\delta}{2+\delta}}\right)^{k-j}\left(\frac{2}{\lambda_T}\cdot\right)^2 \|m(\xi_j)-m(x)\|_{2+\delta}^2\Bigg]\\
\leq\;& \frac{1}{4\pi^2 T_2^2 M^2}\left(\left[\frac{4T_2 M}{\lambda_T^2 c_T^2}+\frac{4T_2 M(M-1)}{\lambda_T^2}\right]\mathrm{E}\left((m(\xi_j)-m(x))^2\right)\right.\\
&\qquad \left. +2M^2 T_2 4\frac{2+\delta}{\delta}\frac{1}{1-\rho_\xi^{\frac{\delta}{2+\delta}}}\frac{4}{\lambda_T^2}\|m(\xi_j)-m(x)\|_{2+\delta}^2\right)\\
=\;& \frac{1}{\pi^2 T_2 \lambda_T^2}\left(\left[\frac{1}{M c_T^2}+\frac{M-1}{M}\right]\mathrm{E}\left((m(\xi_j)-m(x))^2\right)\right.\\
&\qquad \left. +8\frac{2+\delta}{\delta}\frac{1}{1-\rho_\xi^{\frac{\delta}{2+\delta}}}\|m(\xi_j)-m(x)\|_{2+\delta}^2\right)
\end{aligned}
$$

□

**A.16 Lemma.**
*Let A1, A2, A4, A5, A6 and A7 be fulfilled*

*a) If $T$ is large enough, such that*

$$\lambda_T^2\left(m'(x)\left\|f_\xi'\right\|_\infty+\frac{1}{2}\|m''\|_\infty f_\xi(x)\right)+\lambda_T^3\frac{1}{2}\|m''\|_\infty\left\|f_\xi'\right\|_\infty \mathrm{E}\left(|U|^3\right)\leq \epsilon/2$$

*then*

$$
\begin{aligned}
\mathrm{P}\left(|J_T(x)|>\epsilon\right) \;\leq\;& \frac{4}{\epsilon^2\pi^2 T_2\lambda_T^2}\left(\left[\frac{1}{M c_T^2}+\frac{M-1}{M}\right]\mathrm{E}\left((m(\xi_j)-m(x))^2\right)\right.\\
&\qquad \left. +8\frac{2+\delta}{\delta}\frac{1}{1-\rho_\xi^{\frac{\delta}{2+\delta}}}\|m(\xi_j)-m(x)\|_{2+\delta}^2\right)
\end{aligned}
$$

*b) If additionally all tuning parameters are chosen as in* (2.17) *and in* (2.21), *then*

$$J_T(x) \xrightarrow[T\to\infty]{p} 0.$$

**Proof:**

*a)*

$$\begin{aligned}
\mathrm{P}\left(\left|J_T\right|(x) > \epsilon\right) &\leq \mathrm{P}\left(\left|J_T(x) - \mathrm{E}\left(J_T(x)\right)\right| > \epsilon - \left|\mathrm{E}\left(J_T(x)\right)\right|\right) \\
&\leq \mathrm{P}\left(\left|J_T\right|(x) > \epsilon\right) \leq \mathrm{P}\left(\left|J_T(x) - \mathrm{E}\left(J_T(x)\right)\right| > \epsilon/2\right) \\
&\leq \frac{4}{\epsilon^2}\mathrm{Var}\left(J_T(x)\right) \\
&\leq \frac{4}{\epsilon^2\pi^2 T_2 \lambda_T^2}\left(\left[\frac{1}{Mc_T^2} + \frac{M-1}{M}\right]\mathrm{E}\left((m(\xi_j) - m(x))^2\right)\right. \\
&\qquad\left. + 8\frac{2+\delta}{\delta}\frac{1}{1-\rho_\xi^{\frac{\delta}{2+\delta}}}\left\|m(\xi_j) - m(x)\right\|_{2+\delta}^2\right)
\end{aligned}$$

*where the third inequality is due to Lemma A.15a) and the fact, that $T$ is chosen large enough and the last inequality is due to Lemma A.15b).*

*b) a) implies*

$$\begin{aligned}
\mathrm{P}\left(\left|J_T\right|(x) > \epsilon\right) &\leq \mathrm{P}\left(\left|J_T(x) - \mathrm{E}\left(J_T(x)\right)\right| > \epsilon - \left|\mathrm{E}\left(J_T(x)\right)\right|\right) \\
&= O(T^{2c_2c-1}(\log(T))^2) \stackrel{T\to\infty}{\longrightarrow} 0,
\end{aligned}$$

*since $2c_2c - 1 < 0$.*

□

# Appendix B

# Proofs of Chapter 3

The next three Lemmas will be necessary in the following calculations:

**B.1 Lemma.**
*If $X$ is $\alpha$-mixing, $Y \in L^\infty(\sigma(X_s, s \leq t))$ and $Z \in L^\infty(\sigma(X_s, s \geq t+k))$ then the following Covariance-inequality is fulfilled*

$$|\mathrm{Cov}\,(Y, Z)| \leq 4 \, \|Y\|_\infty \, \|Z\|_\infty \, \alpha(k).$$

**Proof:**
*The Lemma is an immediate consequence of Corollary 1.1 in Bosq (1996) [21] with $q = r = \infty$ and $p = 1$.*

□

**B.2 Lemma.**
*Let $f$, $g$ be real valued functions and $X, Y, U, V$ be four real-valued random-variables, where $U$ and $V$ are independent and the random-vector $(U, V)$ is independent from $X$ and $Y$. Then*

$$\mathrm{Cov}\,(f(X+U), g(Y+V)) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{Cov}\,(f(X+u), g(Y+v)) \, dP^U(u) \, dP^V(v).$$

**Proof:**

$$\begin{aligned}
&\mathrm{Cov}\,(f(X+U), g(Y+V)) \\
&\quad = \mathrm{E}\,(f(X+U) g(Y+V)) - \mathrm{E}\,(f(X+U))\, \mathrm{E}\,(g(Y+V)) \\
&\quad = \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}^2} f(x+u) g(y+v) \, dP^{(X,Y)}(x,y) \, dP^U(u) \, dP^V(v) \\
&\qquad - \int_{\mathbb{R}} \int_{\mathbb{R}} f(x+u) \, dP^X(x) \, dP^U(u) \int_{\mathbb{R}} \int_{\mathbb{R}} g(y+v) \, dP^Y(y) \, dP^V(v)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{I\!R}\int_{I\!R} \left[\mathrm{E}\left(f(X+u)g(Y+v)\right) - \mathrm{E}\left(f(X+u)\right)\mathrm{E}\left(g(Y+v)\right)\right] \, dP^U(u)\, dP^V(v) \\
&= \int_{I\!R}\int_{I\!R} \mathrm{Cov}\left(f(X+u), g(Y+v)\right) dP^U(u)\, dP^V(v)
\end{aligned}
$$

□

**B.3 Lemma.**
*Let $X \in L^6(P)$, $c > 0$ and A11 be fulfilled, then:*

*a)*

$$
\int_{I\!R}\int_{I\!R} K(s)\left(s + \frac{t}{c}\right)^2 ds\, dP^X(t) \;=\; \mathrm{E}\left(U^2\right) + \frac{1}{c^2}\mathrm{E}\left(X^2\right),
$$

*b)*

$$
\begin{aligned}
\int_{I\!R}\int_{I\!R} K(s)\left|s + \frac{t}{c}\right|^3 ds\, dP^X(t) \;\leq\;& \mathrm{E}\left(|U|^3\right) + 2\mathrm{E}\left(U^2\right)\frac{1}{c}\sqrt{\mathrm{E}\left(X\right)^2} \\
&+ 2\mathrm{E}\left(|U|\right)\frac{1}{c^2}\mathrm{E}\left(X^2\right) + \frac{1}{c^3}\mathrm{E}\left(X^4\right) + \frac{1}{c^3}\mathrm{E}\left(X^2\right),
\end{aligned}
$$

*c)*

$$
\int_{I\!R}\int_{I\!R} K(s)\left(s + \frac{t}{c}\right)^4 ds\, dP^X(t) \;=\; \mathrm{E}\left(U^4\right) + \frac{1}{c^2}\mathrm{E}\left(U^2\right)\mathrm{E}\left(X^2\right),
$$

*d)*

$$
\begin{aligned}
\int_{I\!R}\int_{I\!R} K(s)\left|s + \frac{t}{c}\right|^5 ds\, dP^X(t) \;\leq\;& \mathrm{E}\left(|U|^5\right) + 5\frac{1}{c}\mathrm{E}\left(U^4\right)\sqrt{\mathrm{E}\left(X^2\right)} \\
&+ 10\frac{1}{c^2}\mathrm{E}\left(|U|^3\right)\mathrm{E}\left(X^2\right) + 10\frac{1}{c^3}\mathrm{E}\left(U^2\right)\mathrm{E}\left(X^4\right) \\
&+ 10\frac{1}{c^3}\mathrm{E}\left(U^2\right)\mathrm{E}\left(X^2\right) + 5\frac{1}{c^4}\mathrm{E}\left(|U|\right)\mathrm{E}\left(X^4\right) \\
&+ \frac{1}{c^5}\mathrm{E}\left(X^6\right) + \frac{1}{c^5}\mathrm{E}\left(X^4\right).
\end{aligned}
$$

**Proof:**
*In the following calculations these estimations will be needed:*

$$
\begin{aligned}
E\left(|X|^k\right) &\leq E\left(X^{k+1}\right) + E\left(X^{k-1}\right), \qquad X \in L^{k+1}(P),\ k \text{ odd}, \\
E\left(|X|\right) &\leq \sqrt{E\left(X^2\right)}, \qquad X \in L^2(P).
\end{aligned}
$$

*a)*

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)\left(s+\frac{t}{c}\right)^2 ds\, dP^X(t)\\
= &\int_{\mathbb{R}} \mathrm{E}\left(\left(U+\frac{t}{c}\right)^2\right) dP^X(t) = \mathrm{E}\left(U^2\right) + \frac{1}{c^2}\mathrm{E}\left((X)^2\right)
\end{aligned}$$

*b)*

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)\left|s+\frac{t}{c}\right|^3 ds\, dP^X(t)\\
\leq\ &\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)|s|^3 ds\, dP^X(t) + 2\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)s^2\frac{|t|^2}{c}\, ds\, dP^X(t)\\
&+2\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)|s|\frac{t^2}{c^2}\, ds\, dP^X(t) + \int_{\mathbb{R}}\int_{\mathbb{R}} K(s)\frac{|t|^3}{c^3}\, ds\, dP^X(t)\\
=\ &\mathrm{E}\left(|U|^3\right) + 2\mathrm{E}\left(U^2\right)\frac{1}{c}\mathrm{E}\left(|X|\right) + 2\mathrm{E}\left(|U|\right)\frac{1}{c^2}\mathrm{E}\left((X)^2\right) + \frac{1}{c^3}\mathrm{E}\left(|X|^3\right)\\
\leq\ &\mathrm{E}\left(|U|^3\right) + 2\mathrm{E}\left(U^2\right)\frac{1}{c}\sqrt{\mathrm{E}\left(X\right)^2} + 2\mathrm{E}\left(|U|\right)\frac{1}{c^2}\mathrm{E}\left(X^2\right) + \frac{1}{c^3}\mathrm{E}\left(X^4\right) + \frac{1}{c^3}\mathrm{E}\left(X^2\right)
\end{aligned}$$

*c)*

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)\left(s+\frac{t}{c}\right)^4 ds\, dP^X(t)\\
=\ &\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)s^4 ds\, dP^X(t) + 3\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)s^3\frac{t}{c}\, ds\, dP^X(t)\\
&+6\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)s^2\frac{t^2}{c^2}\, ds\, dP^X(t) + 3\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)s\frac{t^3}{c^3}\, ds\, dP^X(t)\\
&+\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)\frac{t^4}{c^4}\, ds\, dP^X(t)\\
=\ &\mathrm{E}\left(U^4\right) + \frac{1}{c^2}\mathrm{E}\left(U^2\right)\mathrm{E}\left(X^2\right)
\end{aligned}$$

*d)*

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)\left|s+\frac{t}{c}\right|^5 ds\, dP^X(t)\\
\leq\ &\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)\left(|s|^5 + 5s^4\frac{|t|}{c} + 10|s|^3\frac{t^2}{c^2} + 10s^2\frac{|t|^3}{c^3} + 5s\frac{t^4}{c^4} + \frac{t^5}{|c|^5}\right) ds\, dP^X(t)
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{E}\left(|U|^5\right) + 5\frac{1}{c}\mathrm{E}\left(U^4\right)\mathrm{E}\left(|X|\right) + 10\frac{1}{c^2}\mathrm{E}\left(|U|^3\right)\mathrm{E}\left((X)^2\right) \\
&\quad + 10\frac{1}{c^3}\mathrm{E}\left(U^2\right)\mathrm{E}\left(|X|^3\right) + 5\frac{1}{c^4}\mathrm{E}\left(|U|\right)\mathrm{E}\left((X)^4\right) + \frac{1}{c^5}\mathrm{E}\left(|X|^5\right) \\
&\leq \mathrm{E}\left(|U|^5\right) + 5\frac{1}{c}\mathrm{E}\left(U^4\right)\sqrt{\mathrm{E}\left(X^2\right)} + 10\frac{1}{c^2}\mathrm{E}\left(|U|^3\right)\mathrm{E}\left(X^2\right) \\
&\quad + 10\frac{1}{c^3}\mathrm{E}\left(U^2\right)\mathrm{E}\left(X^4\right) + 10\frac{1}{c^3}\mathrm{E}\left(U^2\right)\mathrm{E}\left(X^2\right) + 5\frac{1}{c^4}\mathrm{E}\left(|U|\right)\mathrm{E}\left(X^4\right) \\
&\quad + \frac{1}{c^5}\mathrm{E}\left(X^6\right) + \frac{1}{c^5}\mathrm{E}\left(X^4\right)
\end{aligned}
$$

□

## B.1 Proof of Theorem 3.4

Theorem 3.4 is a direct implication of the following Lemma:

**B.4 Lemma.**
*If A1, A2, A8, A10, A11, A12 and A13b are fulfilled, then*

*a)*

$$\mathrm{E}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\right) = f_\xi(x) + O\left(T^{-0.4}\right) + O\left(M_T^{-1}\right)$$

*b)*

$$\mathrm{Var}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\right) = O\left(T^{-0.6}\right) + O\left(T^{-0.6}M_T^{-1}\right)$$

**Proof:**

*a)*

$$
\begin{aligned}
&E\left(\frac{1}{T\lambda_T}\sum_{t=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\right) \\
=\quad & \frac{1}{\lambda_T}E\left(K\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{\lambda_T}\int_{\mathbb{R}}\int_{\mathbb{R}} K\left(\frac{x-u-v}{\lambda_T}\right) f_\xi(u)\,du\,dP^{\epsilon^T}(v)\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right) f_\xi(x-\lambda_T s)\,ds\,dP^{\epsilon^T}(t)\\
&= \int_{\mathbb{R}}\left[\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right) f_\xi(x)\,ds - \int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right)\lambda_T s f_\xi'(x)\,ds\right.\\
&\qquad\left.+\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right)\frac{1}{2}\lambda_T^2 s^2 f_\xi''(\hat{x})ds\right] dP^{\epsilon^T}(t)\\
&= \int_{\mathbb{R}} f_\xi(x)\,dP^{\epsilon^T} - f_\xi'(x)\lambda_T\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)\left(s+\frac{t}{\lambda_T}\right) ds\,dP^{\epsilon^T}(t) + C_{T,x}\\
&\overset{K\ symm}{=} f_\xi(x) - f_\xi'(x)\int_{\mathbb{R}} t\,dP^{\epsilon^T}(t) + C_{T,x} = f_\xi(x) - f_\xi'(x)\mathrm{E}\left(\epsilon_1^T\right) + C_{T,x}\\
&= f_\xi(x) + C_{T,x},
\end{aligned}
$$

*where $\hat{x}$ denotes an appropriate value between $x$ and $(x-\lambda_T s)$ and is therefore dependent from $s$. $C_{T,x}$ is defined as follows:*

$$
C_{T,x} := -\int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right)\frac{1}{2}\lambda_T^2 s^2 f_\xi''(\hat{x})\,ds\,dP^{\epsilon^T}(t)
$$

*and $|C_{T,x}|$ is bounded in the following way:*

$$
\begin{aligned}
|C_{T,x}| &= \left|\int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right)\frac{1}{2}\lambda_T^2 s^2 f_\xi''(\hat{x})\,ds\,dP^{\epsilon^T}(t)\right|\\
&\le \frac{1}{2}\lambda_T^2\left\|f_\xi''\right\|_\infty\left|\int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right) s^2 ds\,dP^{\epsilon^T}(t)\right|\\
&= \frac{1}{2}\lambda_T^2\left\|f_\xi''\right\|_\infty\left|\int_{\mathbb{R}}\int_{\mathbb{R}} K(s)\left(s+\frac{t}{\lambda_T}\right)^2 ds\,dP^{\epsilon^T}(t)\right|\\
&\overset{L.\ B.3a)}{\le} \frac{1}{2}\mathrm{E}\left(U^2\right)\lambda_T^2\left\|f_\xi''\right\|_\infty + \frac{1}{2}\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\left\|f_\xi''\right\|_\infty\\
&= O\left(T^{-0,4}\right) + O\left(M_T^{-1}\right).
\end{aligned}
$$

*b)*

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{1}{T\lambda_T}\sum_{t=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\right)\\
&\quad= \frac{1}{T\lambda_T^2}\mathrm{Var}\left(K\left(\frac{x-\hat{\xi}_1^T}{\lambda_T}\right)\right)
\end{aligned}
$$

$$+\frac{1}{T^2\lambda_T^2}\sum_{\substack{i,j=1\\ i\neq j}}^{T}\mathrm{Cov}\left(K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right),K\left(\frac{x-\hat{\xi}_j^T}{\lambda_T}\right)\right),$$

*where*

$$\begin{aligned}
&\mathrm{Var}\left(K\left(\frac{x-\hat{\xi}_1^T}{\lambda_T}\right)\right)\\
&= \mathrm{Var}\left(K\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)\right) \le \mathrm{E}\left(K^2\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)\right)\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} K^2\left(\frac{x-u-v}{\lambda_T}\right) f_\xi(u)\,du\,dP^{\epsilon^T}(v)\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} K^2\left(s-\frac{t}{\lambda_T}\right) f_\xi(x-\lambda_T s)\,ds\,dP^{\epsilon^T}(t)\\
&\le \int_{\mathbb{R}}\int_{\mathbb{R}} \|K\|_\infty\, K\left(s-\frac{t}{\lambda_T}\right) f_\xi(x-\lambda_T s)\,ds\,dP^{\epsilon^T}(t)\\
&= \|K\|_\infty \int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right) f_\xi(x-\lambda_T s)\,ds\,dP^{\epsilon^T}(t)\\
&\le \|K\|_\infty\left(f_\xi(x)+\frac{\mathrm{E}\,(U^2)}{2}\lambda_T^2\left\|f_\xi''\right\|_\infty+\frac{1}{2}\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\left\|f_\xi''\right\|_\infty\right),
\end{aligned}$$

*where the last inequality is just shown in a). For $i\neq j$ we get:*

$$\begin{aligned}
&\left|\mathrm{Cov}\left(K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right),K\left(\frac{x-\hat{\xi}_j^T}{\lambda_T}\right)\right)\right|\\
&\le \int_{\mathbb{R}^2}\left|\mathrm{Cov}\left(K\left(\frac{x-\xi_i-u}{\lambda_T}\right),K\left(\frac{x-\xi_j-v}{\lambda_T}\right)\right)\right| d\left(P^{\epsilon^T}\right)^2(u,v)\\
&\le 4\,\|K\|_\infty^2\,\alpha_\xi(|i-j|) = 4\,\|K\|_\infty^2\,\varrho_\xi^{|i-j|},
\end{aligned}$$

*where the first inequality is due to Lemma B.2, the second to Corollary 1.1 in Bosq (1996) [21] and the third to the fact that the chain is geometrical $\alpha$-mixing as ensured by A1 and A2. As described in 1.1 Assumptions $\varrho_\xi$ is an appropriate real value smaller 1 with $\alpha_\xi(k)\le\varrho_\xi^k$.*
*Thus we achieve:*

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{T\lambda_T}\sum_{t=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\right)\\
&\le \frac{\|K\|_\infty}{T\lambda_T^2}f_\xi(x)+\frac{\mathrm{E}\,(U^2)\,\|K\|_\infty}{2T}\left\|f_\xi''\right\|_\infty+\frac{\|K\|_\infty}{2T\lambda_T^2}\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\left\|f_\xi''\right\|_\infty\\
&\quad+\frac{1}{T^2\lambda_T^2}\sum_{\substack{i,j=1\\ i\neq j}}^{T}4\,\|K\|_\infty^2\,\varrho_\xi^{|i-j|}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\|K\|_\infty}{T\lambda_T^2} f_\xi(x) + \frac{\mathrm{E}\left(U^2\right)\|K\|_\infty}{2T}\left\|f_\xi''\right\|_\infty + \frac{\|K\|_\infty}{2T\lambda_T^2}\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\left\|f_\xi''\right\|_\infty \\
&\quad + \frac{8\|K\|_\infty^2}{T^2\lambda_T^2}\sum_{i=1}^T\sum_{j=i+1}^T \varrho_\xi^{j-i} \\
&= \frac{\|K\|_\infty}{T\lambda_T^2} f_\xi(x) + \frac{\mathrm{E}\left(U^2\right)\|K\|_\infty}{2T}\left\|f_\xi''\right\|_\infty + \frac{\|K\|_\infty}{2T\lambda_T^2}\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\left\|f_\xi''\right\|_\infty \\
&\quad + \frac{8\|K\|_\infty^2}{T^2\lambda_T^2}\sum_{i=1}^T \frac{1}{1-\varrho_\xi} \\
&= \frac{\|K\|_\infty}{T\lambda_T^2} f_\xi(x) + \frac{\mathrm{E}\left(U^2\right)\|K\|_\infty}{2T}\left\|f_\xi''\right\|_\infty + \frac{\|K\|_\infty}{2T\lambda_T^2}\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\left\|f_\xi''\right\|_\infty \\
&\quad + \frac{8(1-\varrho_\xi)^{-1}\|K\|_\infty^2}{T\lambda_T^2} \\
&= O\left(T^{-0.6}\right) + O\left(T^{-1}\right) + O\left(T^{-0.6}M_T^{-1}\right) + O\left(T^{-0.6}\right) \\
&= O\left(T^{-0.6}\right) + O\left(T^{-0.6}M_T^{-1}\right).
\end{aligned}$$

□

## B.2 Proof of Theorem 3.2

Since we already know that the denominator of (3.5) converges to $f_\xi(x)$, the proof of Theorem 3.2 directly follows from the following Proposition, Theorem 3.4 and Slutzki's Theorem

**B.5 Proposition.**
*If A1, A2, A3, A8, A9, A10, A11, A12 and A13a are fulfilled, then*

$$\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^T K\left(\frac{x-\hat{\xi}_i}{\lambda_T}\right)\left(\hat{\xi}_{i+1} - m(x)\right) \xrightarrow[T\to\infty]{d} \mathcal{N}\left(c_\lambda \mathrm{E}\left(U^2\right)\left(-m'(x)f_\xi'(x) + \frac{1}{2}f_\xi(x)m''(x)\right), f_\xi(x)\sigma_\eta^2\|K\|_{L_2}^2\right)$$

Since we have to establish some auxiliary results first, the proof of Proposition

B.5 will be postponed until the end of this section.
Let us divide the numerator into three parts, namely

$$\begin{aligned}
&\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i}{\lambda_T}\right)\left(\hat{\xi}_{i+1}-m(x)\right)\\
&= \frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i}{\lambda_T}\right)\eta_{i+1} \qquad \text{(B.1)}\\
&\quad + \frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i}{\lambda_T}\right)\epsilon_{i+1}^T + \frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i}{\lambda_T}\right)(m(\xi_i)-m(x))
\end{aligned}$$

and study the limiting behaviour of these three parts:

**B.6 Lemma.**
*If A1, A2, A8, A9, A10, A11 and A12 are fulfilled, then*

*a)*

$$\begin{aligned}
&\mathrm{E}\left(\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)(m(\xi_i)-m(x))\right)\\
&= \mathrm{E}\left(U^2\right)\left(-m'(x)f_\xi'(x)+\frac{1}{2}f_\xi(x)m''(x)\right)\sqrt{T\lambda_T^5}\\
&\quad +O\left(T^{-0.2}\right)+O\left(T^{0.4}M_T^{-1}\right)+O\left(M_T^{-0.5}\right),
\end{aligned}$$

*b)*

$$\mathrm{Var}\left(\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)(m(\xi_i)-m(x))\right) = O\left(T^{-0.2}\right)+\left(M_T^{-0.5}\right).$$

**Proof:**

*a)*

$$\begin{aligned}
&\mathrm{E}\left(\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)(m(\xi_i)-m(x))\right)\\
&= \sqrt{\frac{T}{\lambda_T}}\mathrm{E}\left(K\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)(m(\xi_1)-m(x))\right)
\end{aligned}$$

$$
\begin{aligned}
&= \sqrt{\frac{T}{\lambda_T}} \int_{\mathbb{R}} \int_{\mathbb{R}} K\left(\frac{x-u-v}{\lambda_T}\right) (m(u)-m(x))\, f_\xi(u)\, du\, dP^{\epsilon^T}(v) \\
&= \sqrt{T\lambda_T} \int_{\mathbb{R}} \int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right) (m(x-\lambda_T s)-m(x))\, f_\xi(x-\lambda_T s)\, ds\, dP^{\epsilon^T}(t) \\
&= \sqrt{T\lambda_T} \int_{\mathbb{R}} \int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right) \left(-\lambda_T s m'(x) + \frac{1}{2}\lambda_T^2 s^2 m''(x) - \frac{1}{6}\lambda_T^3 s^3 m'''(\hat{x})\right) \\
&\qquad \cdot \left(f_\xi(x) - \lambda_T s f_\xi'(x) + \frac{1}{2}\lambda_T^2 s^2 f_\xi''\left(\hat{\hat{x}}\right)\right) ds\, dP^{\epsilon^T}(t) \\
&= (i) + (ii) + (iii) + (iv) + (v),
\end{aligned}
$$

*with $\hat{x}$ and $\hat{\hat{x}}$ denoting appropriate values between $x$ and $(x-\lambda_T s)$ and*

$$
\begin{aligned}
(i) &= -\sqrt{T\lambda_T}\lambda_T \int_{\mathbb{R}} \int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right) s\, m'(x) f_\xi(x)\, ds\, dP^{\epsilon^T}(t) \\
&= -\sqrt{T\lambda_T}\lambda_T m'(x) f_\xi(x) \int_{\mathbb{R}} \int_{\mathbb{R}} K(s) \left(s+\frac{t}{\lambda_T}\right) ds\, dP^{\epsilon^T}(t) \\
&\overset{K\ symm}{=} -\sqrt{T\lambda_T}\lambda_T m'(x) f_\xi(x) \int_{\mathbb{R}} \frac{t}{\lambda_T} dP^{\epsilon^T}(t) \\
&= -\sqrt{T\lambda_T} m'(x) f_\xi(x) \int_{\mathbb{R}} t dP^{\epsilon^T}(t) \\
&= -\sqrt{T\lambda_T} m'(x) f_\xi(x) \mathrm{E}\left(\epsilon_1^T\right) = 0,
\end{aligned}
$$

$$
\begin{aligned}
(ii) &= \sqrt{T\lambda_T} \\
&\quad \cdot \int_{\mathbb{R}^2} K\left(s-\frac{t}{\lambda_T}\right) \lambda_T^2 s^2 \left(-m'(x) f_\xi'(x) + \frac{1}{2} f_\xi(x) m''(x)\right) ds\, dP^{\epsilon^T}(t) \\
&= \sqrt{T\lambda_T}\lambda_T^2 \left(-m'(x) f_\xi'(x) + \frac{1}{2} f_\xi(x) m''(x)\right) \\
&\quad \cdot \int_{\mathbb{R}} \int_{\mathbb{R}} K(s) \left(s+\frac{t}{\lambda_T}\right)^2 ds\, dP^{\epsilon^T}(t) \\
&\overset{L.\ B.3a)}{=} \sqrt{T\lambda_T}\lambda_T^2 \left(-m'(x) f_\xi'(x) + \frac{1}{2} f_\xi(x) m''(x)\right) \mathrm{E}\left(U^2\right) \\
&\quad + \sqrt{T\lambda_T} \left(-m'(x) f_\xi'(x) + \frac{1}{2} f_\xi(x) m''(x)\right) \mathrm{E}\left(\left(\epsilon_1^T\right)^2\right) \\
&= \sqrt{T\lambda_T^5} \left(-m'(x) f_\xi'(x) + \frac{1}{2} f_\xi(x) m''(x)\right) \mathrm{E}\left(U^2\right) + O\left(T^{0.4} M_T^{-1}\right),
\end{aligned}
$$

$$
\begin{aligned}
|(iii)| &= \left| \sqrt{T\lambda_T} \int_{\mathbb{R}} \int_{\mathbb{R}} K(s) \lambda_T^3 \left(s+\frac{t}{\lambda_T}\right)^3 \right. \\
&\quad \left. \cdot \left(-\frac{1}{2} m'(x) f_\xi''\left(\hat{\hat{x}}\right) + \frac{1}{2} m''(x) f_\xi'(x) - \frac{1}{3} m'''(\hat{x})\, f_\xi(x)\right) ds\, dP^{\epsilon^T}(t) \right|
\end{aligned}
$$

$$
\begin{aligned}
&\le \frac{1}{2}\sqrt{T\lambda_T}\lambda_T^3 \int_{I\!R}\int_{I\!R} K(s)\left|s+\frac{t}{\lambda_T}\right|^3 ds\, dP^{\epsilon^T}(t) \\
&\quad \cdot\left(|m'(x)|\left\|f_\xi''\right\|_\infty + |m''(x)|\left|f_\xi'(x)\right| + \frac{1}{3}\left\|m'''\right\|_\infty |f_\xi(x)|\right) \\
&\overset{L.\ B.3b)}{\le} \frac{1}{2}\sqrt{T\lambda_T}\lambda_T^3 \\
&\quad \left(|m'(x)|\left\|f_\xi''\right\|_\infty + |m''(x)|\left|f_\xi'(x)\right| + \frac{1}{3}\left\|m'''\right\|_\infty |f_\xi(x)|\right) \\
&\quad \cdot\left(\mathrm{E}\left(|U|^3\right) + \frac{2}{\lambda_T}\mathrm{E}\left(U^2\right)\sqrt{\mathrm{E}\left((\epsilon_1^T)^2\right)} + 2\mathrm{E}\left(|U|\right)\frac{1}{\lambda_T^2}\mathrm{E}\left((\epsilon_1^T)^2\right)\right. \\
&\quad \left. + \frac{1}{\lambda_T^3}\left[\mathrm{E}\left((\epsilon_1^T)^4\right) + \mathrm{E}\left((\epsilon_1^T)^2\right)\right]\right) \\
&= O\left(T^{-0.2}\right) + O\left(M_T^{-0.5}\right) + O\left(T^{0.2}M_T^{-1}\right) + O\left(T^{0.4}M_T^{-1}\right) \\
&= O\left(T^{-0.2}\right) + O\left(M_T^{-0.5}\right) + O\left(T^{0.4}M_T^{-1}\right),
\end{aligned}
$$

$$
\begin{aligned}
|(iv)| &= \left|\frac{1}{2}\sqrt{T\lambda_T}\int_{I\!R}\int_{I\!R} K\left(s-\frac{t}{\lambda_T}\right)\lambda_T^4 s^4\right. \\
&\qquad \left. \cdot\left(\frac{1}{2}m''(x)f_\xi''\left(\hat{\hat{x}}\right) - \frac{1}{3}m'''\left(\hat{x}\right)f_\xi'(x)\right) ds\, dP^{\epsilon^T}(t)\right| \\
&\le \frac{1}{2}\sqrt{T\lambda_T}\lambda_T^4\left(\frac{1}{2}m''(x)\left\|f_\xi''\right\|_\infty + \frac{1}{3}\left\|m'''\right\|_\infty\left|f_\xi'(x)\right|\right) \\
&\quad \cdot\int_{I\!R}\int_{I\!R} K(s)\left(s+\frac{t}{\lambda_T}\right)^4 ds\, dP^{\epsilon^T}(t) \\
&\overset{L.\ B.3c)}{=} \frac{1}{2}\sqrt{T\lambda_T}\lambda_T^4\left(\frac{1}{2}m''(x)\left\|f_\xi''\right\|_\infty + \frac{1}{3}\left\|m'''\right\|_\infty\left|f_\xi'(x)\right|\right) \\
&\quad \cdot\left(\mathrm{E}\left(U^4\right) + \frac{1}{\lambda_t^2}\mathrm{E}\left(U^2\right)\mathrm{E}\left((\epsilon_1^T)^2\right)\right) \\
&= O\left(T^{-0.4}\right) + O\left(M_T^{-1}\right),
\end{aligned}
$$

$$
\begin{aligned}
|(v)| &= \left|\frac{1}{12}\sqrt{T\lambda_T}\int_{I\!R}\int_{I\!R} K\left(s-\frac{t}{\lambda_T}\right)\lambda_T^5 s^5\left(-m'''\left(\hat{x}\right)f_\xi''\left(\hat{\hat{x}}\right)\right) ds\, dP^{\epsilon^T}(t)\right| \\
&\le \frac{1}{12}\sqrt{T\lambda_T}\lambda_T^5\left\|m'''\right\|_\infty\left\|f_\xi''\right\|_\infty\int_{I\!R}\int_{I\!R} K(s)\left|s+\frac{t}{\lambda_T}\right|^5 ds\, dP^{\epsilon^T}(t) \\
&\overset{L.\ B.3d)}{\le} \frac{1}{12}\sqrt{T\lambda_T}\lambda_T^5\left\|m'''\right\|_\infty\left\|f_\xi''\right\|_\infty \\
&\quad \left(\mathrm{E}\left(|U|^5\right) + \frac{5}{\lambda_T}\mathrm{E}\left(U^4\right)\sqrt{\mathrm{E}\left((\epsilon_1^T)^2\right)} + \frac{10}{\lambda_T^2}\mathrm{E}\left(|U|^3\right)\mathrm{E}\left((\epsilon_1^T)^2\right)\right. \\
&\quad \left. + \frac{10}{\lambda_T^3}\mathrm{E}\left(U^2\right)\left[\mathrm{E}\left((\epsilon_1^T)^4\right) + \mathrm{E}\left((\epsilon_1^T)^2\right)\right] + \frac{5}{\lambda_T^4}\mathrm{E}\left(|U|\right)\mathrm{E}\left((\epsilon_1^T)^4\right)\right.
\end{aligned}
$$

$$
\begin{aligned}
& \left. +\frac{1}{\lambda_T^5}\left[\mathrm{E}\left((\epsilon_1^T)^6\right)+\mathrm{E}\left((\epsilon_1^T)^4\right)\right]\right) \\
= \quad & O\left(T^{-0.6}\right)+O\left(T^{-0.4}M_T^{-0.5}\right)+O\left(T^{0.4}M_T^{-1}\right).
\end{aligned}
$$

*Putting (i) - (v) together a) is proved.*

*b)*

$$
\begin{aligned}
& \mathrm{Var}\left(\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)(m(\xi_i)-m(x))\right) \\
= \quad & \sum_{i=1}^{T}\mathrm{Var}\left(K\left(\frac{x-\xi_i-\epsilon_i^T}{\lambda_T}\right)(m(\xi_i)-m(x))\right) \\
& +\sum_{\substack{i,j=1\\ i\neq j}}^{T}\mathrm{Cov}\left(K\left(\frac{x-\xi_i-\epsilon_i^T}{\lambda_T}\right)(m(\xi_i)-m(x)), K\left(\frac{x-\xi_j-\epsilon_j^T}{\lambda_T}\right)(m(\xi_j)-m(x))\right),
\end{aligned}
$$

*where*

$$
\begin{aligned}
& \mathrm{Var}\left(K\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)(m(\xi_1)-m(x))\right) \\
\leq \quad & E\left(K^2\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)(m(\xi_1)-m(x))^2\right) \\
= \quad & \int_{\mathbb{R}}\int_{\mathbb{R}} K^2\left(\frac{x-u-v}{\lambda_T}\right)(m(u)-m(x))^2 f_\xi(u)\,du\,dP^{\epsilon^T}(v) \\
= \quad & \lambda_T\int_{\mathbb{R}}\int_{\mathbb{R}} K^2\left(s-\frac{t}{\lambda_T}\right)(m(x-\lambda_T s)-m(x))^2 f_\xi(x-\lambda_t s)\,ds\,dP^{\epsilon^T}(t) \\
\leq \quad & \lambda_T\int_{\mathbb{R}}\int_{\mathbb{R}} \|K\|_\infty\, K\left(s-\frac{t}{\lambda_T}\right)\left(-\lambda_T s\, m'\left(\hat{x}\right)\right)^2 \\
& \qquad \cdot\left(f_\xi(x)-\lambda_T s f_\xi'(x)+\frac{1}{2}\lambda_T^2 s^2 f_\xi''(\hat{\hat{x}})\right)ds\,dP^{\epsilon^T}(t) \\
\leq \quad & \lambda_T^3\|K\|_\infty\|m'\|_\infty \\
& \cdot\left[\int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right)s^2 f_\xi(x)\,ds\,dP^{\epsilon^T}(t)\right. \\
& \quad +\int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right)|s|^3\lambda_T\left|f_\xi'(x)\right|ds\,dP^{\epsilon^T}(t) \\
& \quad \left.+\int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s-\frac{t}{\lambda_T}\right)s^4\lambda_T^2\left\|f_\xi''\right\|_\infty ds\,dP^{\epsilon^T}(t)\right] \\
\leq \quad & \lambda_T^3\|K\|_\infty\|m'\|_\infty \\
& \cdot\left\{f_\xi(x)\left[\mathrm{E}\left(U^2\right)+\frac{1}{\lambda_T^2}\mathrm{E}\left((\epsilon^T)^2\right)\right]\right.
\end{aligned}
$$

$$
\begin{aligned}
&+\lambda_T \left|f_\xi(x)\right| \Bigg[\mathrm{E}\left(|U|^3\right) + \frac{3}{\lambda_T}\mathrm{E}\left(U^2\right)\sqrt{\mathrm{E}\left((\epsilon^T)^2\right)} + \frac{3}{\lambda_T^2}\mathrm{E}\left(|U|\right)\mathrm{E}\left((\epsilon^T)^2\right) \\
&\qquad + \frac{1}{\lambda_T^3}\left(\mathrm{E}\left((\epsilon^T)^4\right) + \mathrm{E}\left((\epsilon^T)^2\right)\right)\Bigg] \\
&+\lambda_T^2 \left\|f_\xi''\right\|_\infty \left[\mathrm{E}\left(U^4\right) + \frac{6}{\lambda_T^2}\mathrm{E}\left(U^2\right)\mathrm{E}\left((\epsilon^T)^2\right) + \frac{1}{\lambda_T^4}\mathrm{E}\left((\epsilon^T)^4\right)\right]\Bigg\} \\
= \;& O\left(T^{-0.6}\right) + O\left(T^{-0.2}M_T^{-1}\right) + O\left(T^{-0.8}\right) + O\left(T^{-0.6}M_T^{-0.5}\right) \\
&+O\left(T^{-0.4}M_T^{-1}\right) + O\left(T^{-0.2}M_T^{-1}\right) + O\left(T^{-1}\right) + O\left(T^{-0.6}M_T^{-1}\right) \\
&+O\left(T^{-0.2}M_T^{-1}\right) \\
= \;& O\left(T^{-0.6}\right) + O\left(T^{-0.2}M_T^{-1}\right) + O\left(T^{-0.6}M_T^{-0.5}\right)
\end{aligned}
$$

*and for $i \neq j$ we achieve by using Corollary 1.1 of Bosq (1996)*

$$
\begin{aligned}
&\left|\mathrm{Cov}\left(K\left(\frac{x-\xi_i-\epsilon_i^T}{\lambda_T}\right)(m(\xi_i)-m(x)), K\left(\frac{x-\xi_j-\epsilon_j^T}{\lambda_T}\right)(m(\xi_j)-m(x))\right)\right| \\
\overset{L.\ B.2}{\leq}\;& \int_{\mathbb{R}^2}\Bigg|\mathrm{Cov}\Bigg(K\left(\frac{x-\xi_i-u}{\lambda_T}\right)(m(\xi_i)-m(x)), \\
&\qquad K\left(\frac{x-\xi_j-v}{\lambda_T}\right)(m(\xi_j)-m(x))\Bigg)\Bigg|\, d\left(P^{\epsilon^T}\right)^2(u,v) \\
\leq\;& \int_{\mathbb{R}^2} 4 \sup_{x,y\in\mathbb{R}}\left\{K\left(\frac{x-y-u}{\lambda_T}\right)(m(y)-m(x))\right\} \\
&\qquad \cdot \sup_{x,y\in\mathbb{R}}\left\{K\left(\frac{x-y-v}{\lambda_T}\right)(m(y)-m(x))\right\}\alpha_\xi(|i-j|)\, d\left(P^{\epsilon^T}\right)^2(u,v) \\
\leq\;& 4\int_{\mathbb{R}^2}\|K\|_\infty^2 \sup\left\{(y-x)m'(\theta(x,y))\,|-\lambda_T-u \leq y-x \leq -u+\lambda_T\right\} \\
&\qquad \cdot \sup\left\{(y-x)m'(\theta(x,y))\,|-\lambda_T-v \leq y-x \leq -v+\lambda_T\right\} \\
&\qquad \cdot \alpha_\xi(|i-j|)\, d\left(P^{\epsilon^T}\right)^2(u,v) \\
\leq\;& 4\|K\|_\infty^2 \int_{\mathbb{R}^2}(\lambda_T+|u|)(\lambda_T+|v|)\,\|m'\|_\infty^2\,\alpha_\xi(|i-j|)\, d\left(P^{\epsilon^T}\right)^2(u,v) \\
=\;& 4\alpha_\xi(|i-j|)\,\|K\|_\infty^2\,\|m'\|_\infty^2 \\
&\cdot\int_{\mathbb{R}^2}\left(\lambda_T^2+\lambda_T|u|+\lambda_T|v|+|u|\,|v|\right) d\left(P^{\epsilon^T}\right)^2(u,v) \\
=\;& 4\alpha_\xi(|i-j|)\,\|K\|_\infty^2\,\|m'\|_\infty^2\left(\lambda_T^2+2\lambda_T\mathrm{E}\left(\left|\epsilon_1^T\right|\right)+\mathrm{E}\left(\left|\epsilon_1^T\right|\right)^2\right) \\
\leq\;& 4\varrho_\xi^{|i-j|}\,\|K\|_\infty^2\,\|m'\|_\infty^2\left(\lambda_T^2+2\lambda_T\sqrt{\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)}+\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\right).
\end{aligned}
$$

*Thus we get:*

$$
\begin{aligned}
&\operatorname{Var}\left(\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T}K\left(\frac{x-\xi_i-\epsilon_i^T}{\lambda_T}\right)(m(\xi_i)-m(x))\right)\\
\leq\ & \frac{1}{T\lambda_T}\Bigg(T\left[O\left(T^{-0.6}\right)+O\left(T^{-0.2}M_T^{-1}\right)+O\left(T^{-0.6}M_T^{-0.5}\right)\right]\\
&\qquad +4\left\|K\right\|_\infty^2\left\|m'\right\|_\infty^2\left(\lambda_T^2+2\lambda_T\sqrt{\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)}+\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\right)2\sum_{i=1}^{T}\sum_{j=i+1}^{T}\varrho_\xi^{j-i}\Bigg)\\
\leq\ & \frac{1}{\lambda_T}\left[O\left(T^{-0.6}\right)+O\left(T^{-0.2}M_T^{-1}\right)+O\left(T^{-0.6}M_T^{-0.5}\right)\right]\\
&+\frac{8}{\lambda_T}\left\|K\right\|_\infty^2\left\|m'\right\|_\infty^2\left(\lambda_T^2+2\lambda_T\sqrt{\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)}+\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\right)\frac{1}{1-\varrho_\xi}\\
=\ & O\left(T^{-0.4}\right)+O\left(M_T^{-1}\right)+O\left(T^{-0.4}M_T^{-0.5}\right)+O\left(T^{-0.2}\right)+O\left(M_T^{-0.5}\right)+O\left(M_T^{-1}\right)\\
=\ & O\left(T^{-0.2}\right)+O\left(M_T^{-0.5}\right).
\end{aligned}
$$

□

**B.7 Corollary.**
*If A1, A2, A8, A9, A10, A11 and A12 are fulfilled, then*

*a)*

$$
\begin{aligned}
&\mathrm{E}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)(m(\xi_i)-m(x))\right)\\
&\qquad = \mathrm{E}\left(U^2\right)\left(-m'(x)f'_\xi(x)+\frac{1}{2}f_\xi(x)m''(x)\right)\lambda_T^4\\
&\qquad\quad +O\left(T^{-0.6}\right)+\left(M_T^{-1}\right)+O\left(T^{-0.4}M_T^{-0.5}\right),
\end{aligned}
$$

*b)*

$$
\operatorname{Var}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)(m(\xi_i)-m(x))\right)=O\left(T^{-1}\right)+\left(T^{-0.8}M_T^{-0.5}\right).
$$

Lemma B.6 directly implies:

**B.8 Proposition.**
*If A1, A2, A8, A9, A10, A11, A12 and A13a are fulfilled, then*

$$\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)(m(\xi_i)-m(x))$$
$$\xrightarrow[T\to\infty]{p} \quad c_\lambda \mathrm{E}\left(U^2\right)\left(-m'(x)f_\xi(x)+\frac{1}{2}f_\xi'(x)m''(x)\right).$$

**B.9 Lemma.**
*If A1, A2, A8, A10, A11 and A12 are fulfilled, then*

*a)*

$$\mathrm{E}\left(\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\epsilon_{i+1}^T\right)=0,$$

*b)*

$$\mathrm{Var}\left(\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\epsilon_{i+1}^T\right)=O\left(M_T^{-1}\right).$$

**Proof:**

*a)*

$$\begin{aligned}
&\mathrm{E}\left(\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\epsilon_{i+1}^T\right)\\
&\quad = \frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T}\mathrm{E}\left(K\left(\frac{x-\xi_i-\epsilon_i^T}{\lambda_T}\right)\right)\mathrm{E}\left(\epsilon_{i+1}^T\right)\\
&\quad = 0.
\end{aligned}$$

*b)*

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\epsilon_{i+1}^T\right)\\
&= \sum_{i=1}^{T}\mathrm{Var}\left(K\left(\frac{x-\xi_i-\epsilon_i^T}{\lambda_T}\right)\epsilon_{i+1}^T\right)
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{\substack{i,j=1\\ i\neq j}}^{T}\underbrace{\operatorname{Cov}\left(K\left(\frac{x-\xi_i-\epsilon_i^T}{\lambda_T}\right)\epsilon_{i+1}^T, K\left(\frac{x-\xi_j-\epsilon_j^T}{\lambda_T}\right)\epsilon_{j+1}^T\right)}_{=0}\\
= &\; T\left(\mathrm{E}\left(K^2\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)\left(\epsilon_2^T\right)^2\right)-\underbrace{\left(\mathrm{E}\left(K\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)\epsilon_2^T\right)\right)^2}_{=0}\right)\\
= &\; TE\left(K^2\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)\right)E\left(\left(\epsilon_2^T\right)^2\right)\\
\leq &\; T\lambda_T\left\|K\right\|_\infty\left[\frac{1}{\lambda_T}E\left(K\left(\frac{x-\xi_1-\epsilon_1^T}{\lambda_T}\right)\right)\right]\mathrm{E}\left(\left(\epsilon_2^T\right)^2\right)\\
\leq &\; T\lambda_T\left\|K\right\|_\infty\left(f_\xi(x)+\frac{\mathrm{E}\left(U^2\right)}{2}\lambda_T^2\left\|f_\xi''\right\|_\infty+\frac{1}{2}\mathrm{E}\left(\left(\epsilon_1^T\right)^2\right)\left\|f_\xi''\right\|_\infty\right)\mathrm{E}\left(\left(\epsilon_2^T\right)^2\right)\\
= &\; O\left(T^{0.8}\right)O\left(M_T^{-1}\right)+O\left(T^{0.4}\right)O\left(M_T^{-1}\right)+O\left(T^{0.8}\right)O\left(M_T^{-2}\right)\\
= &\; O\left(T^{0.8}\right)O\left(M_T^{-1}\right),
\end{aligned}
$$

*where the last inequality is shown in the proof of Lemma B.4a). Thus we get*

$$
\begin{aligned}
\operatorname{Var}\left(\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\epsilon_{i+1}^T\right) &= \frac{1}{T\lambda_T}O\left(T^{0.8}\right)O\left(M_T^{-1}\right)\\
&= O\left(M_T^{-1}\right).
\end{aligned}
$$

□

**B.10 Corollary.**
*If A1, A2, A8, A10, A11 and A12 are fulfilled, then*

*a)*

$$\mathrm{E}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\epsilon_{i+1}^T\right)=0,$$

*b)*

$$\operatorname{Var}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\epsilon_{i+1}^T\right)=O\left(T^{-0.8}\right)O\left(M_T^{-1}\right).$$

Lemma B.9 directly implies:

**B.11 Proposition.**
*If A1, A2, A8, A10, A11, A12 and A13a are fulfilled, then*

$$\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\epsilon_{i+1}^T \xrightarrow[T\to\infty]{p} 0.$$

We will now show, that the last Summand of $(B.1)$ converges in distribution to a certain normal distribution. Therefore we first show in Lemma B.13

$$\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\xi_i}{\lambda_T}\right)\eta_{i+1} \xrightarrow[T\to\infty]{d} \mathcal{N}\left(0, f_\xi(x)\sigma_\eta^2\,\|K\|_{L_2}^2\right) \tag{B.2}$$

and afterwards in Lemma B.14

$$\mathrm{E}\left(\left[\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T}\left[K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)-K\left(\frac{x-\xi_i}{\lambda_T}\right)\right]\eta_{i+1}\right]^2\right) \stackrel{T\to\infty}{\longrightarrow} 0. \tag{B.3}$$

Proposition B.12 is then an immediate consequence of Lemma B.13 and Lemma B.14:

**B.12 Proposition.**
*If A1, A2, A3, A8, A10, A11, A12 and A13a are fulfilled, then*

$$\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T} K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\eta_{i+1} \xrightarrow[T\to\infty]{d} \mathcal{N}\left(0, f_\xi(x)\sigma_\eta^2\,\|K\|_{L_2}^2\right).$$

**B.13 Lemma.**
*If A1, A2, A3, A10, A11 and A12 are fulfilled, then $(B.2)$ is true.*
**Proof:**
*We will use the following Central Limit Theorem which was introduced in Brown (1971) [22]:*

**Brown's Martingale Central Limit Theorem**
*Let $\{S_n, \mathcal{F}_n, n \in I\!N\}$ be a martingale on the probability space $(\Omega, \mathcal{F}, P)$, with $S_0 = 0$ and $X_n = S_n - S_{n-1}$, $n \in I\!N$. Let*

$$V_n^2 := \sum_{j=1}^{n}\mathrm{E}\left(X_j^2|\mathcal{F}_{j-1}\right) \qquad s_n := \sqrt{E(V_n^2)}.$$

*Then*

$$\frac{S_n}{s_n} \xrightarrow[T\to\infty]{d} \mathcal{N}(0,1)$$

*if*

*(1)* $V_n^2 s_n^{-2} \xrightarrow[n\to\infty]{p} 1$ *and*

*(2)* $s_n^{-2} \sum_{j=1}^n \mathrm{E}\left(X_j^2 1_{\{|X_j|>\epsilon s_n\}} | \mathcal{F}_{j-1}\right) \xrightarrow[n\to\infty]{p} 0 \quad \forall \epsilon > 0.$

*Let* $S_T = \sum_{i=1}^T K\left(\frac{x-\xi_i}{\lambda_T}\right)\eta_{i+1}$, *then* $X_i = K\left(\frac{x-\xi_i}{\lambda_T}\right)\eta_{i+1}$.
*Defining* $\mathcal{F}_T := \sigma\left(\{\xi_j, \eta_j, j \leq T\}\right)$ *it is easy to check, that* $(S_T)_{T\in\mathbb{N}}$ *becomes a martingale w.r.t.* $(\mathcal{F}_T)_{T\in\mathbb{N}}$.

$$\begin{aligned}
\mathrm{E}\left(S_{T+1}|\mathcal{F}_T\right) &= \mathrm{E}\left(S_T + X_T|\mathcal{F}_T\right) = S_T + \mathrm{E}\left(K\left(\frac{x-\xi_T}{\lambda_T}\right)\eta_{T+1}\middle|\mathcal{F}_T\right) \\
&= S_T + K\left(\frac{x-\xi_T}{\lambda_T}\right)\mathrm{E}\left(\eta_{T+1}|\mathcal{F}_T\right) = S_T + K\left(\frac{x-\xi_T}{\lambda_T}\right)\mathrm{E}\left(\eta_{T+1}\right) \\
&= S_T.
\end{aligned}$$

*We now want to check condition (1):*

$$\begin{aligned}
\mathrm{E}\left(V_T^2\right) &= \mathrm{E}\left(\sum_{j=1}^T \mathrm{E}\left(K^2\left(\frac{x-\xi_j}{\lambda_T}\right)\eta_{j+1}^2\middle|\mathcal{F}_{j-1}\right)\right) = \mathrm{E}\left(\sum_{j=1}^T K^2\left(\frac{x-\xi_j}{\lambda_T}\right)\mathrm{E}\left(\eta_{j+1}^2|\mathcal{F}_{j-1}\right)\right) \\
&= \sigma_\eta^2 \sum_{j=1}^T \mathrm{E}\left(K^2\left(\frac{x-\xi_j}{\lambda_T}\right)\right) = T\lambda_T\sigma_\eta^2 \int_{\mathbb{R}} K^2\left(\frac{x-u}{\lambda_T}\right) f_\xi(u)\frac{1}{\lambda_T}du \\
&= T\lambda_T\sigma_\eta^2 \int_{\mathbb{R}} K^2(s)\, f_\xi(x-\lambda_T s)ds \\
&= T\lambda_T\sigma_\eta^2 \int_{\mathbb{R}} K^2(s)\left[f_\xi(x) - \lambda_T s f_\xi(\theta(s))\right]ds \\
&= T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x) + c_{T,x}\right),
\end{aligned}$$

*where* $\theta(s)$ *denotes an appropriate interim value and* $c_{T,x} := -\int_{\mathbb{R}} K^2(s)\lambda_T s f_\xi(\theta(s))\,ds$, *which implies* $|c_{T,x}| \leq \|K\|_{L^2}^2 \left\|f'_\xi\right\|_\infty \lambda_T \overset{T\to\infty}{\longrightarrow} 0$.

$$\begin{aligned}
&\mathrm{Var}\left(V_T^2 s_t^{-2}\right) \\
&= \mathrm{Var}\left(\frac{1}{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x) + c_{T,x}\right)}\sum_{j=1}^T \mathrm{E}\left(K^2\left(\frac{x-\xi_j}{\lambda_T}\right)\eta_{j+1}^2\middle|\mathcal{F}_{j-1}\right)\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{T^2\lambda_{T+1}^2\sigma_\eta^4\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)^2}\mathrm{Var}\left(\sum_{j=1}^{T}K^2\left(\frac{x-\xi_j}{\lambda_T}\right)\mathrm{E}\left(\eta_{j+1}^2|\mathcal{F}_{j-1}\right)\right)\\
&= \frac{1}{T^2\lambda_{T+1}^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)^2}\sum_{i=1}^{T}\mathrm{Var}\left(K^2\left(\frac{x-\xi_i}{\lambda_T}\right)\right)\\
&\quad+\frac{2}{T^2\lambda_{T+1}^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)^2}\sum_{i=1}^{T}\sum_{j=i+1}^{T}\mathrm{Cov}\left(K^2\left(\frac{x-\xi_i}{\lambda_T}\right),K^2\left(\frac{x-\xi_j}{\lambda_T}\right)\right)\\
&\leq \frac{1}{T^2\lambda_{T+1}^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)^2}\sum_{i=1}^{T}\mathrm{E}\left(K^4\left(\frac{x-\xi_i}{\lambda_T}\right)\right)\\
&\quad+\frac{2}{T^2\lambda_{T+1}^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)^2}\sum_{i=1}^{T}\sum_{j=i+1}^{T}4\|K\|_\infty^4\ \alpha_\xi(j-i)\\
&\leq \frac{\|K\|_\infty^2}{T\lambda_{T+1}^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)^2}\mathrm{E}\left(K^2\left(\frac{x-\xi_1}{\lambda_T}\right)\right)\\
&\quad+\frac{8\|K\|_\infty^4}{T^2\lambda_{T+1}^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)^2}\sum_{i=1}^{T}\frac{1}{1-\varrho_\xi}\\
&\leq \frac{\|K\|_\infty^2}{T\lambda_T\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)^2}\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)\\
&\quad+\frac{8\|K\|_\infty^4}{T\lambda_{T+1}^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)^2(1-\varrho_\xi)}\\
&= O(T^{-0,6})\xrightarrow{T\to\infty}0.
\end{aligned}
$$

*Thus*

$$V_T^2 s_T^{-2}\ \underset{T\to\infty}{\xrightarrow{p}}\ 1.$$

*Now we are going to check condition (2). Let $\epsilon,\delta>0$, then:*

$$
\begin{aligned}
&\frac{1}{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\sum_{j=1}^{T}\mathrm{E}\left(X_j^2 1_{\left\{|X_j|>\epsilon\sqrt{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2 f_\xi(x)+c_{T,x}\right)}\right\}}\middle|\mathcal{F}_{j-1}\right)\\
&= \frac{1}{\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\\
&\quad\cdot\mathrm{E}\left(K^2\left(\frac{x-\xi_1}{\lambda_T}\right)\eta_2^2 1_{\left\{\left|K\left(\frac{x-\xi_1}{\lambda_T}\right)\eta_2\right|>\epsilon\sqrt{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2 f_\xi(x)+c_{T,x}\right)}\right\}}\middle|\mathcal{F}_0\right)\\
&\leq \frac{1}{\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}
\end{aligned}
$$

$$\cdot \mathrm{E}\left( K^2\left(\frac{x-\xi_1}{\lambda_T}\right)\eta_2^2 1_{\left\{|\eta_2|>\frac{\epsilon}{\|K\|_\infty}\sqrt{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\right\}}\middle|\mathcal{F}_0\right)$$
$$\leq \frac{1}{\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}K^2\left(\frac{x-\xi_1}{\lambda_T}\right)\mathrm{E}\left(\eta_2^2 1_{\left\{|\eta_2|>\frac{\epsilon}{\|K\|_\infty}\sqrt{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\right\}}\right)$$

*Thus*

$$P\left(\frac{1}{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\sum_{j=1}^{T}\mathrm{E}\left(X_j^2 1_{\left\{|X_j|>\epsilon\sqrt{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\right\}}\middle|\mathcal{F}_{j-1}\right)>\delta\right)$$
$$\leq \frac{1}{\delta\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\mathrm{E}\left(K^2\left(\frac{x-\xi_1}{\lambda_T}\right)\right)$$
$$\cdot \mathrm{E}\left(\eta_2^2 1_{\left\{|\eta_2|>\frac{\epsilon}{\|K\|_\infty}\sqrt{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\right\}}\right)$$
$$= \frac{\lambda_T\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}{\delta\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\mathrm{E}\left(\eta_2^2 1_{\left\{|\eta_2|>\frac{\epsilon}{\|K\|_\infty}\sqrt{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\right\}}\right)$$
$$= \frac{1}{\delta\sigma_\eta^2}\mathrm{E}\left(\eta_2^2 1_{\left\{|\eta_2|>\frac{\epsilon}{\|K\|_\infty}\sqrt{T\lambda_T\sigma_\eta^2\left(\|K\|_{L_2}^2\ f_\xi(x)+c_{T,x}\right)}\right\}}\right),$$

*where the first inequality is just an application of Markov's inequality.*
*Since the set of the indicator-function converges to the empty set for* $T\to\infty$*, the whole expression converges to* $0$ *and (2) is fulfilled.*
*Thus we have*

$$\frac{S_T}{s_T}\xrightarrow[T\to\infty]{d}\mathcal{N}(0,1)$$

*and therefore:*

$$\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T}K\left(\frac{x-\xi_i}{\lambda_T}\right)\eta_{i+1}$$
$$= \frac{S_T}{s_T}\sigma_\eta\|K\|_{L^2}\sqrt{f_\xi(x)}\cdot\left(\frac{\|K\|_{L^2}^2\ f_\xi(x)+c_{T,x}}{\|K\|_{L^2}^2\ f_\xi(x)}\right)^{0,5}$$
$$\xrightarrow[T\to\infty]{d}\mathcal{N}(0,\sigma_\eta^2\|K\|_{L^2}^2\ f_\xi(x)),$$

*because the last factor converges to* $1$ *since* $c_{T,x}$ *converges to* $0$.

□

**B.14 Lemma.**
*If A3, A8, A10, A11, A12 and A13a are fulfilled, then* $(B.3)$ *is true.*
**Proof:**

$$
\begin{aligned}
&\mathrm{E}\left(\left[\frac{1}{\sqrt{T\lambda_T}}\sum_{i=1}^{T}\left[K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)-K\left(\frac{x-\xi_i}{\lambda_T}\right)\right]\eta_{i+1}\right]^2\right)\\
&= \frac{1}{\sqrt{T\lambda_T}}\\
&\quad \sum_{i,j=1}^{T}\underbrace{\mathrm{E}\left(\left[K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)-K\left(\frac{x-\xi_i}{\lambda_T}\right)\right]\left[K\left(\frac{x-\hat{\xi}_j^T}{\lambda_T}\right)-K\left(\frac{x-\xi_j}{\lambda_T}\right)\right]\eta_{i+1}\eta_{j+1}\right)}_{=0 \ \mathit{for}\ i\neq j}\\
&= \frac{\sigma_\eta}{T\lambda_T}\mathrm{E}\left(\sum_{i=1}^{T}\left[K\left(\frac{x-\xi_i-\epsilon_i^T}{\lambda_T}\right)-K\left(\frac{x-\xi_i}{\lambda_T}\right)\right]^2\right)\\
&= \frac{\sigma_\eta}{\lambda_T}\int_{I\!R}\int_{I\!R}\left[K\left(\frac{x-u-v}{\lambda_T}\right)-K\left(\frac{x-u}{\lambda_T}\right)\right]^2 f_\xi(u)\,du\,dP^{\epsilon^T}(v)\\
&= \frac{\sigma_\eta}{\lambda_T}\lambda_T\int_{I\!R}\int_{I\!R}\left[K\left(s-\frac{t}{\lambda_T}\right)-K(s)\right]^2 f_\xi(x-\lambda_T s)\,ds\,dP^{\epsilon^T}(t)\\
&= \sigma_\eta\int_{I\!R}\int_{I\!R}\left[K\left(s-\frac{t}{\lambda_T}\right)-K(s)\right]^2\left[f_\xi(x)-\lambda_T s f_\xi'(x)+\frac{\lambda_T^2 s^2}{2}f_\xi''(\hat{x})\right]ds\,dP^{\epsilon^T}(t)\\
&\leq \underbrace{\sigma_\eta\int_{I\!R}\int_{I\!R}\left[K\left(s-\frac{t}{\lambda_T}\right)-K(s)\right]^2 f_\xi(x)\,ds\,dP^{\epsilon^T}(t)}_{:=S1}\\
&\quad \underbrace{+\sigma_\eta\left|\int_{I\!R}\int_{I\!R}\left[K\left(s-\frac{t}{\lambda_T}\right)-K(s)\right]^2\lambda_T s f_\xi'(x)ds\,dP^{\epsilon^T}(t)\right|}_{:=S2}\\
&\quad \underbrace{+\sigma_\eta\lambda_T^2\left|\int_{I\!R}\int_{I\!R}\left[K\left(s-\frac{t}{\lambda_T}\right)-K(s)\right]^2\frac{s^2}{2}f_\xi''(\hat{x})ds\,dP^{\epsilon^T}(t)\right|}_{:=S3}
\end{aligned}
$$

*Remember* $\mathrm{supp}(K)=[-a,a]\ a\in I\!R$*:*

$$
\begin{aligned}
S1 &= \sigma_\eta\int_{I\!R}\int_{\left\{s:\left[K\left(s-\frac{t}{\lambda_T}\right)-K(s)\right]\neq 0\right\}}\|K'\|^2\ \frac{t^2}{\lambda_T^2}f_\xi(x)\,ds\,dP^{\epsilon^T}(t)\\
&= \sigma_\eta\|K'\|^2\ \frac{1}{\lambda_T^2}f_\xi(x)\int_{I\!R}\int_{\left\{s:\left[K\left(s-\frac{t}{\lambda_T}\right)-K(s)\right]\neq 0\right\}}1\,ds\,t^2 dP^{\epsilon^T}(t)
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sigma_\eta \left\|K'\right\|^2 \frac{1}{\lambda_T^2} f_\xi(x) \cdot 4a \cdot \mathrm{E}\left(\left(\epsilon_1^T\right)^2\right) \\
&= O\left(T^{0.4}\right) \cdot O\left(M_T^{-1}\right) = O\left(\left(\frac{T^2}{M_T^5}\right)^{1/5}\right) \stackrel{T\to\infty}{\longrightarrow} 0,
\end{aligned}
$$

*since A13a is fulfilled.*

$$
\begin{aligned}
S2 &= \sigma_\eta \lambda_T \left|f_\xi'(x)\right| \left| \underbrace{\int_{\mathbb{R}}\int_{\mathbb{R}} K^2\left(s - \frac{t}{\lambda_T}\right) s\, ds\, dP^{\epsilon^T}(t)}_{=:S2a} + \underbrace{\int_{\mathbb{R}}\int_{\mathbb{R}} K^2(s)\, s\, ds\, dP^{\epsilon^T}(t)}_{=:S2b} \right. \\
&\qquad \left. -2\int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s - \frac{t}{\lambda_T}\right) K(s) s\, ds\, dP^{\epsilon^T}(t) \right|
\end{aligned}
$$

$$
\begin{aligned}
S2a &= \int_{\mathbb{R}}\int_{\mathbb{R}} K^2(s)\left(s + \frac{t}{\lambda_T}\right) ds\, dP^{\epsilon^T}(t) \stackrel{K\ symm}{=} \int_{\mathbb{R}}\int_{\mathbb{R}} K^2(s) \frac{t}{\lambda_T}\, ds\, dP^{\epsilon^T}(t) \\
&= \frac{1}{\lambda_T}\int_{\mathbb{R}} \left\|K\right\|_{L^2}^2\, t\, dP^{\epsilon^T}(t) = \frac{\left\|K\right\|_{L^2}^2}{\lambda_T} \mathrm{E}\left(\epsilon_1^T\right) = 0 \\
S2b &= \int_{\mathbb{R}}\int_{\mathbb{R}} K^2(s) s ds\, dP^{\epsilon^T}(t) \stackrel{K\ symm}{=} 0.
\end{aligned}
$$

*Hence:*

$$
\begin{aligned}
S2 &\leq 2\sigma_\eta \lambda_T \left|f_\xi'(x)\right| \int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s - \frac{t}{\lambda_T}\right) K(s) \left|s\right| ds\, dP^{\epsilon^T}(t) \\
&\leq 2\sigma_\eta \lambda_T \left|f_\xi'(x)\right| \int_{\mathbb{R}}\int_{\mathbb{R}} \left\|K\right\|_\infty K(s)\left|s\right| ds\, dP^{\epsilon^T}(t) \\
&= 2\sigma_\eta \lambda_T \left|f_\xi'(x)\right| \left\|K\right\|_\infty \mathrm{E}\left(\left|U\right|\right) \\
&= O\left(T^{-0.2}\right)
\end{aligned}
$$

$$
\begin{aligned}
S3 &\leq \sigma_\eta \lambda_T^2 \left\|f_\xi''\right\|_\infty \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}} \left[K\left(s - \frac{t}{\lambda_T}\right) - K(s)\right]^2 s^2 ds\, dP^{\epsilon^T}(t) \\
&= \sigma_\eta \lambda_T^2 \left\|f_\xi''\right\|_\infty \frac{1}{2} \left[ \underbrace{\int_{\mathbb{R}}\int_{\mathbb{R}} K^2\left(s - \frac{t}{\lambda_T}\right) s^2 ds\, dP^{\epsilon^T}(t)}_{=:S3a} + \underbrace{\int_{\mathbb{R}}\int_{\mathbb{R}} K^2(s)\, s^2 ds\, dP^{\epsilon^T}(t)}_{=:S3b} \right. \\
&\qquad \left. + \underbrace{(-2)\int_{\mathbb{R}}\int_{\mathbb{R}} K\left(s - \frac{t}{\lambda_T}\right) K(s)\, s^2 ds\, dP^{\epsilon^T}(t)}_{=:S3c} \right].
\end{aligned}
$$

*By using Lemma B.3 we achieve*

$$
\begin{aligned}
S3a &\leq \|K\|_\infty \int_{\mathbb{R}} \int_{\mathbb{R}} K(s) \left(s + \frac{t}{\lambda_T}\right)^2 ds\, dP^{\epsilon^T}(t) \\
&= \|K\|_\infty \operatorname{E}\left(U^2\right) + \|K\|_\infty \frac{\operatorname{E}\left((\epsilon_1^T)^2\right)}{\lambda_T^2}, \\
S3b &\leq \|K\|_\infty \operatorname{E}\left(U^2\right), \\
|S3c| &\leq 2 \int_{\mathbb{R}} \int_{\mathbb{R}} \|K\|_\infty K(s) s^2 ds\, dP^{\epsilon^T}(t) \leq 2 \|K\|_\infty \operatorname{E}\left(U^2\right).
\end{aligned}
$$

*So we get*

$$
\begin{aligned}
S3 &\leq \sigma_\eta \lambda_T^2 \left\|f_\xi''\right\|_\infty \|K\|_\infty \left(4\operatorname{E}\left(U^2\right) + \frac{\operatorname{E}\left((\epsilon_1^T)^2\right)}{\lambda_T^2}\right) \\
&= O\left(T^{-0.4}\right) + O\left(M_T^{-1}\right) \overset{T\to\infty}{\longrightarrow} 0,
\end{aligned}
$$

*which - together with the convergence of* $S1$ *and* $S2$ *to* $0$ *- proves the statement.*

□

**B.15 Proof of Proposition B.5.**

*Since the second and third summand of* $(B.1)$ *converge to constants in probability, as shown in Proposition B.8 and in Proposition B.11, and therefore also in distribution, Slutzki's Theorem says, that the left-side-expression has the same limit as the first summand plus these constants. Together with Proposition B.12 this proves the statement.*

□

## B.3 Proof of Theorem 3.3

We use a decomposition like the one in $(B.1)$ and show $(\hat{m}_T(x) - m(x)) \xrightarrow[T\to\infty]{p} 0$:

$$\begin{aligned} &\hat{m}_T(x) - m(x) \\ &= \hat{f}_\xi(x)^{-1} \left( \frac{1}{T\lambda_T} \sum_{i=1}^{T} K\left(\frac{x - \hat{\xi}_i^T}{\lambda_T}\right) [m(\xi_i) - m(x)] \right. \\ &\quad \left. + \frac{1}{T\lambda_T} \sum_{i=1}^{T} K\left(\frac{x - \hat{\xi}_i^T}{\lambda_T}\right) \eta_{i+1} + \frac{1}{T\lambda_T} \sum_{i=1}^{T} K\left(\frac{x - \hat{\xi}_i^T}{\lambda_T}\right) \epsilon_{i+1}^T \right) \end{aligned} \tag{B.4}$$

From Proposition 3.4 we know that the inverse of the first factor converges to $f(x) \neq 0$. The stochastic convergence to zero of the first and the third addend of the second factor follows rather immediately from the Corollarys B.7 and B.10, respectively.
Corollary B.7 directly implies

**B.16 Proposition.**
*If A1, A2, A9, A10, A11, A12 and A13b are fulfilled, then*

$$\frac{1}{T\lambda_T} \sum_{i=1}^{T} K\left(\frac{x - \hat{\xi}_i^T}{\lambda_T}\right) (m(\xi_i) - m(x)) \xrightarrow[T\to\infty]{p} 0.$$

Corollary B.10 directly implies:

**B.17 Proposition.**
*If A1, A2, A8, A10, A11, A12 and A13b are fulfilled, then*

$$\frac{1}{T\lambda_T} \sum_{i=1}^{T} K\left(\frac{x - \hat{\xi}_i^T}{\lambda_T}\right) \epsilon_{i+1}^T \xrightarrow[T\to\infty]{p} 0.$$

Thus it is just left to show, that the second addend of the second factor of (B.4) converges to zero as well, which is a direct implication of the following Lemma:

**B.18 Lemma.**
*If A1, A2, A3, A8, A10, A11 and A12 are fulfilled, then*

*a)*

$$\mathrm{E}\left( \frac{1}{T\lambda_T} \sum_{i=1}^{T} K\left(\frac{x - \hat{\xi}_i^T}{\lambda_T}\right) \eta_{i+1} \right) = 0,$$

*b)*

$$\mathrm{Var}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\eta_{i+1}\right)=O\left(T^{-0.8}\right).$$

**Proof:**

*a)*

$$\begin{aligned}\mathrm{E}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\eta_{i+1}\right) &= \frac{1}{\lambda_T}\mathrm{E}\left(K\left(\frac{x-\hat{\xi}_1^T}{\lambda_T}\right)\right)\mathrm{E}\left(\eta_2\right)\\ &= 0\end{aligned}$$

*b)*

$$\begin{aligned}&\mathrm{Var}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\eta_{i+1}\right)\\
&= \frac{1}{T^2\lambda_T^2}\mathrm{E}\left(\left(\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\eta_{i+1}\right)^2\right)\\
&= \frac{1}{T\lambda_T^2}\mathrm{E}\left(K^2\left(\frac{x-\hat{\xi}_1^T}{\lambda_T}\right)\eta_2^2\right)\\
&\quad +\frac{2}{T^2\lambda_T^2}\sum_{i=1}^{T}\sum_{j=i+1}^{T}\mathrm{E}\left(K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)K\left(\frac{x-\hat{\xi}_j^T}{\lambda_T}\right)\eta_{i+1}\right)\mathrm{E}\left(\eta_{j+1}\right)\\
&\leq \frac{\sigma_\eta^2}{T\lambda_T}\left\|K\right\|_\infty \mathrm{E}\left(\frac{1}{\lambda_T}K\left(\frac{x-\hat{\xi}_1^T}{\lambda_T}\right)\right)\\
&\leq \frac{\sigma_\eta^2}{T\lambda_T}\left\|K\right\|_\infty \mathrm{E}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T}K\left(\frac{x-\hat{\xi}_i^T}{\lambda_T}\right)\right)\\
&\overset{L.\ B.4a)}{=} \frac{\sigma_\eta^2}{T\lambda_T}\left\|K\right\|_\infty\left(f_\xi(x)+O\left(T^{-0.4}\right)+O\left(M_T^{-1}\right)\right)\\
&= O\left(T^{-0.8}\right).\end{aligned}$$

$\square$

# Appendix C

# Proofs of Chapter 7

**Proof of Lemma 7.2:**
Following Proposition 5.2.1 of Straumann (2005) [94] the stochastic recursive equation

$$s_{t+1} = \psi_t(s_t), \tag{C.1}$$

with

$$\psi_t(x) = m(x^{1/2} \cdot e_t, x).$$

admits a unique stationary solution, if for an arbitrary $\zeta_0^2 \in [0, \infty)$ the following two conditions are fulfilled:

a) $\mathrm{E}\left(\log^+ |\psi_0(\zeta_0^2)|\right) < \infty$.

b) $\mathrm{E}\left(\log^+ \mathrm{Lip}\,(\psi_0)\right) < \infty$ and for some integer $r \geq 1$ it holds, that $\mathrm{E}\left(\log \mathrm{Lip}\,(\psi_0 \circ \ldots \circ \psi_{1-r})\right) < 0$.

If $\sigma_T^2$ is a stationary solution of (C.1), $(X_t = \sigma_t e_t, \sigma_t^2)$ is a stationary solution of (7.1) and (7.2). So let us first check condition $a$):

$$\begin{aligned} \log^+ \left|\psi_0(\zeta_0^2)\right| &= \log^+ \left(m(\zeta_0 e_0, \zeta_0^2)\right) \\ &= m_1(\zeta_0^2 e_0^2, \zeta_0^2) 1_{(-\infty,0)}(e_0) + m_2(\zeta_0^2 e_0^2, \zeta_0^2, \zeta_0^2) 1_{[0,\infty)}(e_0). \end{aligned}$$

Let $e_0 < 0$ and fix an arbitrary $c < 0$

$$\begin{aligned} \left|m_1(\zeta_0^2 e_0^2, \zeta_0^2, \zeta_0^2)\right| &\leq \left|m_1(\zeta_0^2 e_0^2, \zeta_0^2, \zeta_0^2) - m_1(\zeta_0^2 c^2, \zeta_0^2, \zeta_0^2)\right| + \left|m_1(\zeta_0^2 c^2, \zeta_0^2, \zeta_0^2)\right| \\ &\leq C_{\zeta_0,c,1} L_{1,1} \zeta_o^2 \left|(e_0^2 - c^2)\right| + C_{\zeta_0,c,2}, \end{aligned}$$

with $C_{\zeta_0,c,1}$ and $C_{\zeta_0,c,2}$ suitable.
Analogously you get for $e_0 > 0$ and arbitrary $d > 0$

$$\left|m_2(\zeta_0^2 e_0^2, \zeta_0^2, \zeta_0^2)\right| \leq C_{\zeta_0,d,1} L_{1,2} \zeta_o^2 \left|(e_0^2 - d^2)\right| + C_{\zeta_0,d,2},$$

with $C_{\zeta_0,d,1}$ and $C_{\zeta_0,d,2}$ suitable.
Thus

$$\begin{aligned}
\mathrm{E}\left(\log^+\left|\psi_0(\zeta_0^2)\right|\right) &\leq \mathrm{E}\left(\left|\psi_0(\zeta_0^2)\right|\right)\\
&\leq C_{\zeta_0,c,1}L_{1,1}\zeta_0^2\mathrm{E}\left(\left|e_0^2-c^2\right|1_{(-\infty,0)}(e_0)\right)+C_{\zeta_0,c,1}\\
&\quad +C_{\zeta_0,d,1}L_{1,2}\zeta_0^2\mathrm{E}\left(\left|e_0^2-d^2\right|1_{[0,\infty)}(e_0)\right)+C_{\zeta_0,d,2}\\
&\leq C_{\zeta_0,c,1}L_{1,1}\zeta_0^2\mathrm{E}\left((e_0^2+c^2)1_{(-\infty,0)}(e_0)\right)+C_{\zeta_0,c,1}\\
&\quad +C_{\zeta_0,d,1}L_{1,2}\zeta_0^2\mathrm{E}\left((e_0^2+d^2)1_{[0,\infty)}(e_0)\right)+C_{\zeta_0,d,2}\\
&\leq C_1+C_2\sigma_e^2<\infty,
\end{aligned}$$

$C_1$ and $C_2$ suitable.
If the second condition in b) is fulfilled for $r=1$ it is obvious, that the first condition is fulfilled, too. So let us now check, that $\mathrm{E}\left(\log \mathrm{Lip}\left(\psi_0\right)\right)<0$ :

$$\mathrm{Lip}\left(\psi_0\right)=\sup_{x\neq y}\frac{|\psi_0(x)-\psi_0(y)|}{|x-y|}\leq (L_{1,i}e_0^2+L_{2,i}),$$

for i=1,2. Thus by Jensen's inequality we have

$$\begin{aligned}
\mathrm{E}\left(\log\left(\mathrm{Lip}\left(\psi_0\right)\right)\right) &\leq \log\left(\max_{i=1,2}\mathrm{E}\left(L_{1,i}e_0^2+L_{2,i}\right)\right)\\
&= \log\left(\max_{i=1,2}\left(L_{1,i}\sigma_e^2+L_{2,i}\right)\right)\leq \log(L)<0
\end{aligned}$$

Thus (C.1) posesses a stationary solution, namely $(\sigma_t^2)_{t\in\mathbb{Z}}$ and so $X_t=\sigma_t e_t$, $t\in\mathbb{Z}$, is also stationary.
Alternatively one could also use Theorem 3.1 of Straumann and Mikosch (2006) [95], with $C_1(x):=\max_{i=1,2}|L_{1,i}x^2+L_{2,i}|$ and $p=1$ to prove the Lemma. $\square$

**Proof of Lemma 7.4:**
The proof of this Lemma follows the same strategy as the one of Lemma 2.1 of Neumann and Paparoditis (2006) [81].
Since $X_t$ is stationary we assume for simplicity, that $t_u=0$ and denote $(t_1,\ldots,t_{u-1})$ by $(t_1,\ldots,t_n)$.
Let $(\sigma_{t_1},e_{t_1},\ldots,\sigma_{t_n},e_{t_n},\sigma_0,e_0)$ and $(\tilde{\sigma}_{t_1},\tilde{e}_{t_1},\ldots,\tilde{\sigma}_{t_n},\tilde{e}_{t_n},\tilde{\sigma}_0,\tilde{e}_0)$, $t_1<t_2<\ldots<t_n<0$, be independent with the (2n+2)-dimensional stationary distribution of the nonparametric GARCH(1,1) process and $M_0$ be the set of these $(4n+4)$ random-variables. We feed both processes with the same sequence of i.i.d. innovations $(e_t)_{t\in\mathbb{N}}$. Thus $X_t$ and $\tilde{X}_t$ have the same sign if $t>0$. Notice, that if $x$ and $\tilde{x}$ have the same sign

$$|m(x,y)-m(\tilde{x},\tilde{y})|\leq L_{1,i}\left|x^2-\tilde{x}^2\right|+L_{2,i}\left|y-\tilde{y}\right|,$$

with $i = 1, 2$, if $x$ and $\tilde{x}$ are positive or negative, respectively.
Let $t > 1$, $Z_t := \left|X_t^2 - \tilde{X}_t^2\right|$ and $Y_t := |\sigma_t^2 - \tilde{\sigma}_t^2|$. Then

$$\begin{aligned} Z_t &= \left|\sigma_t^2 e_t^2 - \tilde{\sigma}_t^2 e_t^2\right| = \left|m(X_{t-1}, \sigma_{t-1}^2) - m(\tilde{X}_{t-1}, \tilde{\sigma}_{t-1}^2)\right| e_t^2 \\ &\leq \left(L_{1,i}\left|X_{t-1}^2 - \tilde{X}_{t-1}^2\right| + L_{2,i}\left|\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2\right|\right) e_t^2 \\ &= \left(L_{1,i}\left|\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2\right| e_{t-1}^2 + L_{2,i}\left|\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2\right|\right) e_t^2 \\ &= \left|\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2\right| (L_{1,i} e_{t-1}^2 + L_{2,i}) e_t^2 \\ &= Y_{t-1}(L_{1,i} e_{t-1}^2 + L_{2,i}) e_t^2 \end{aligned}$$

and

$$\begin{aligned} Y_t &= \left|m(X_{t-1}, \sigma_{t-1}^2) - m(\tilde{X}_{t-1}, \tilde{\sigma}_{t-1}^2)\right| \leq \left|\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2\right| (L_{1,i} e_{t-1}^2 + L_{2,i}) \\ &= Y_{t-1}(L_{1,i} e_{t-1}^2 + L_{2,i}). \end{aligned}$$

Thus

$$\begin{aligned} \mathrm{E}\left(Z_t | M_0\right) &\leq \max_{i=1,2}\left\{\mathrm{E}\left(Y_{t-1}(L_{1,i} e_{t-1}^2 + L_{2,i}) e_t^2 | M_0\right)\right\} \\ &= \max_{i=1,2}\left\{\mathrm{E}\left(Y_{t-1} | M_0\right)(L_{1,i}\sigma_e^2 + L_{2,i})\sigma_e^2\right\} \\ &\leq \mathrm{E}\left(Y_{t-1} | M_0\right) L\sigma_e^2 \end{aligned}$$

and

$$\mathrm{E}\left(Y_t | M_0\right) \leq L \cdot \mathrm{E}\left(Y_{t-1} | M_0\right).$$

Thus

$$\mathrm{E}\left(Z_t | M_0\right) \leq \mathrm{E}\left(Y_1 | M_0\right) L^{t-1}\sigma_e^2 = \left|m(X_0, \sigma_0^2) - m(\tilde{X}_0, \tilde{\sigma}_0^2)\right| L^{t-1} s_e^2,$$

since $\left|m(X_0, \sigma_0^2) - m(\tilde{X}_0, \tilde{\sigma}_0^2)\right|$ is measureable w.r.t. $M_0$.
Let $P^0$ denote the stationary distribution of $(\sigma_{t_1}, e_{t_1}, \ldots, \sigma_{t_n}, e_{t_n}, \sigma_0, e_0)$ and $\left(X_t^{(s,e)}\right)_{t\in\mathbb{Z}}$ denote the process with known initial value $X_0^{(s,e)} = se$. Then

$$\begin{aligned} &\left|\mathrm{Cov}\left(g\left(X_{t_1}, \ldots, X_{t_n}, X_0\right), h\left(X_{s_1}, \ldots, X_{s_v}\right)\right)\right| \\ &= \mathrm{E}\left(g\left(X_{t_1}, \ldots, X_{t_n}, X_0\right) h\left(X_{s_1}, \ldots, X_{s_v}\right)\right) \\ &\quad -\mathrm{E}\left(g\left(X_{t_1}, \ldots, X_{t_n}, X_0\right)\right)\mathrm{E}\left(h\left(X_{s_1}, \ldots, X_{s_v}\right)\right) \\ &= \Bigg|\int g\left(s_{t_1}e_{t_1}, \ldots, s_{t_n}e_{t_n}, se\right)\mathrm{E}\left(h\left(X_{s_1}^{(s,e)}, \ldots, X_{s_v}^{(s,e)}\right)\right) P^0 d\left(s_{t_1}, e_{t_1}, \ldots, s_{t_n}, e_{t_n}, s, e\right) \\ &\quad - \int g\left(s_{t_1}e_{t_1}, \ldots, s_{t_n}e_{t_n}, se\right) P^0 d\left(s_{t_1}, e_{t_1}, \ldots, s_{t_n}, e_{t_n}, s, e\right) \end{aligned}$$

$$
\begin{aligned}
& \cdot \int \mathrm{E}\left(h\left(X_{s_1}^{(\tilde{s},\tilde{e})},\ldots,X_{s_v}^{(\tilde{s},\tilde{e})}\right)\right) P^0 d\left(\tilde{s}_{t_1},\tilde{e}_{t_1},\ldots,\tilde{s}_{t_n},\tilde{e}_{t_n},\tilde{s},\tilde{e}\right)\Bigg| \\
= \quad & \Bigg|\int\int g\left(s_{t_1}e_{t_1},\ldots,s_{t_n}e_{t_n},se\right)\mathrm{E}\left(h\left(X_{s_1}^{(s,e)},\ldots,X_{s_v}^{(s,e)}\right)-h\left(X_{s_1}^{(\tilde{s},\tilde{e})},\ldots,X_{s_v}^{(\tilde{s},\tilde{e})}\right)\right) \\
& \qquad P^0\otimes P^0 d\left(\left(\tilde{s}_{t_1},\tilde{e}_{t_1},\ldots,\tilde{s}_{t_n},\tilde{e}_{t_n},\tilde{s},\tilde{e}\right),\left(s_{t_1},e_{t_1},\ldots,s_{t_n},e_{t_n},s,e\right)\right)\Bigg| \\
\leq \quad & \sum_{i=1}^{v}\mathrm{Lip}\left(h\right)\int\int\left|g\left(s_{t_1}e_{t_1},\ldots,s_{t_n}e_{t_n},se\right)\right|\mathrm{E}\left(\left|X_{s_i}^{(s,e)}-X_{s_i}^{(\tilde{s},\tilde{e})}\right|\right) \\
& \qquad\qquad P^0\otimes P^0 d\left(\left(s_{t_1},e_{t_1},\ldots,s_{t_n},e_{t_n},s,e\right),\left(\tilde{s}_{t_1},\tilde{e}_{t_1},\ldots,\tilde{s}_{t_n},\tilde{e}_{t_n},\tilde{s},\tilde{e}\right)\right) \\
= \quad & \mathrm{Lip}\left(h\right)\sum_{i=1}^{v}\left(\left[\int\int\left|g\left(s_{t_1}e_{t_1},\ldots,s_{t_n}e_{t_n},se\right)\right|\mathrm{E}\left(\left|X_{s_i}^{(s,e)}-X_{s_i}^{(\tilde{s},\tilde{e})}\right|\right)\right.\right. \\
& \qquad\qquad \left.\left. P^0\otimes P^0 d\left(\left(\ldots\right),\left(\ldots\right)\right)\right]^2\right)^{1/2} \\
\stackrel{\text{C.S.}}{\leq} \quad & \mathrm{Lip}\left(h\right)\sum_{i=1}^{v}\sqrt{\int\int g^2\left(s_{t_1}e_{t_1},\ldots,s_{t_n}e_{t_n},se\right)P^0\otimes P^0 d\left(\left(\ldots\right),\left(\ldots\right)\right)} \\
& \qquad \cdot\sqrt{\int\int\mathrm{E}\left(\left|X_{s_i}^{(s,e)}-X_{s_i}^{(\tilde{s},\tilde{e})}\right|\right)^2 P^0\otimes P^0 d\left(\left(\ldots\right),\left(\ldots\right)\right)} \\
\stackrel{\text{Jensen}}{\leq} \quad & \mathrm{Lip}\left(h\right)\sum_{i=1}^{v}\left(\mathrm{E}\left(g^2\left(X_{t_1},\ldots,X_{t_n},X_0\right)\right)\right)^{1/2} \\
& \cdot\sqrt{\int\int\mathrm{E}\left(\left|X_{s_i}^{(s,e)2}-X_{s_i}^{(\tilde{s},\tilde{e})2}\right|\right)P^0\otimes P^0 d\left(\left(s_{t_1},\ldots,e_{t_n},s,e\right),\left(\tilde{s}_{t_1},\ldots,\tilde{e}_{t_n},\tilde{s},\tilde{e}\right)\right)} \\
\leq \quad & \mathrm{Lip}\left(h\right)\left(\mathrm{E}\left(g^2\left(X_{t_1},\ldots,X_{t_n},X_0\right)\right)\right)^{1/2} \\
& \cdot\sum_{i=1}^{v}\sqrt{\int\int L^{s_{i-1}}\sigma_e^2\left|m\left(se,s^2\right)-m\left(\tilde{s}\tilde{e},\tilde{s}^2\right)\right|P^0\otimes P^0 d\left(\left(\ldots,s,e\right)\right),\left(\ldots,\tilde{s},\tilde{e}\right)} \\
\leq \quad & \mathrm{Lip}\left(h\right)\left(\mathrm{E}\left(g^2\left(X_{t_1},\ldots,X_{t_n},X_0\right)\right)\right)^{1/2} v(L^{(1/2)})^{s_1}L^{-(1/2)}\sigma_e \\
& \cdot\left(\mathrm{E}\left(\left|m\left(X_0,\sigma_0^2\right)-m\left(\tilde{X}_0,\tilde{\sigma}_0^2\right)\right|\right)\right)^{(1/2)} \\
= \quad & C\mathrm{Lip}\left(h\right)\rho_X^{s_1}\left(\mathrm{E}\left(g^2\left(X_{t_1},\ldots,X_{t_n},X_0\right)\right)\right)^{1/2},
\end{aligned}
$$

with suitable constants $C \in I\!R$ and $\rho_X \in [0,1)$.

$\square$

**Proof of Theorem 7.5:**

$$
\begin{aligned}
&\sqrt{T\lambda_T}\left(\hat{f}_X^T(x) - f_X(x)\right) \\
&= \sqrt{T\lambda_T}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T} K\left(\frac{x - X_i}{\lambda_T}\right) - f_X(x)\right) \\
&= \left(\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\frac{1}{\lambda_T}\left[K\left(\frac{x - X_i}{\lambda_T}\right) - \mathrm{E}\left(K\left(\frac{x - X_i}{\lambda_T}\right)\right)\right]\right) \\
&\quad + \sqrt{T\lambda_T}\left(\mathrm{E}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T} K\left(\frac{x - X_i}{\lambda_T}\right)\right) - f_X(x)\right)
\end{aligned}
$$

Let us first show, that the second summand converges to zero, as T goes to infinity:

$$
\begin{aligned}
\mathrm{E}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T} K\left(\frac{x - X_i}{\lambda_T}\right)\right) &= \int \frac{1}{\lambda_T} K\left(\frac{x - u}{\lambda_T}\right) f_X(u)\, du \\
&= \int K(s) f_X(x - \lambda_T s)\, ds \\
&= \int K(s)\left(f_X(x) - \lambda_T s f_X'(x) + \frac{1}{2}\lambda_T^2 s^2 f_X''(\theta)\right) ds \\
&= f_X(x) + C_{T,x},
\end{aligned}
$$

with

$$
|C_{T,x}| = \left|\int \frac{1}{2}\lambda_T^2 s^2 f_X''(\theta)\, ds\right| \leq \frac{1}{2}\lambda_T^2 \left\|f_X''\right\|_\infty \mathrm{E}\left(U^2\right) = O(1)\lambda_T^2.
$$

Thus

$$
\begin{aligned}
\left|\sqrt{T\lambda_T}\left(\mathrm{E}\left(\frac{1}{T\lambda_T}\sum_{i=1}^{T} K\left(\frac{x - X_i}{\lambda_T}\right)\right) - f_X(x)\right)\right| &= \sqrt{T\lambda_T}\,|C_{T,x}| \\
&\leq O(1)\sqrt{T\lambda_T^5} \overset{T\to\infty}{\longrightarrow} 0.
\end{aligned}
$$

Now it remains to show, that the first summand has the desired asymptotic normal distribution. To show this, we make use of the central limit theorem 6.1 formulated in Neumann and Paparoditis (2006) [81]:

> **Central Limit Theorem**
> Suppose that $(X_{T,k})_{k\in\mathbb{Z}}$, $T \in I\!N$, is a triangular scheme of (row-wise) stationary random variables with $\mathrm{E}\left(X_{T,k}\right) = 0$ and $\mathrm{E}\left(X_{T,k}^2\right) \leq C < \infty$. Furthermore, we assume that
>
> $$\frac{1}{T}\sum_{k=1}^{T}\mathrm{E}\left(X_{T,k} 1_{\{|X_{T,k}|/\sqrt{T}>\epsilon\}}\right) \overset{T\to\infty}{\longrightarrow} 0 \qquad (7.5.1)$$

holds for all $\epsilon > 0$ and that

$$\sigma_T^2 = \sum_{k \in \mathbb{Z}} \mathrm{E}\left(X_{T,0} X_{T,k}\right) \stackrel{T \to \infty}{\longrightarrow} \sigma^2 \in [0, \infty). \tag{7.5.2}$$

For $T \geq T_0$, there exists a monotonously nonincreasing and summable sequence $(\theta_r)_{r \in \mathbb{N}}$ such that, for all indices $t_1 < t_2 \ldots < t_u < t_u + r = s_1 \leq s_2$, the following upper bounds for covariances hold true:

- for all measurable and quadratic integrable functions $f : \mathbb{R}^u \to \mathbb{R}$,

$$\begin{aligned} &\left|\mathrm{Cov}\left(f\left(X_{T,t_1}, \ldots, X_{T,t_u}\right), X_{T,s_1}\right)\right| \\ &\qquad \leq \sqrt{\mathrm{E}\left(f^2\left(X_{T,t_1}, \ldots, X_{T,t_u}\right)\right)} \theta_r \end{aligned} \tag{7.5.3}$$

- for all measurable and bounded functions $f : \mathbb{R}^u \to \mathbb{R}$,

$$\left|\mathrm{Cov}\left(f\left(X_{T,t_1}, \ldots, X_{T,t_u}\right), X_{T,s_1} X_{T,s_2}\right)\right| \leq \|f\|_\infty \theta_r \tag{7.5.4}$$

Then

$$\frac{1}{\sqrt{T}}\left(X_{T,1} + \cdots + X_{T,T}\right) \underset{T \to \infty}{\stackrel{d}{\longrightarrow}} \mathcal{N}(0, \sigma^2).$$

Define g: $\mathbb{R} \to \mathbb{R}$ by

$$g_T(y) := \frac{1}{\sqrt{\lambda_T}} K\left(\frac{x - y}{\lambda_T}\right)$$

and

$$\begin{aligned} X_{T,i} &:= g_T(X_i) - \mathrm{E}\left(g_T(X_1)\right) = \frac{1}{\sqrt{\lambda_T}}\left[K\left(\frac{x - X_i}{\lambda_T}\right) - \mathrm{E}\left(K\left(\frac{x - X_1}{\lambda_T}\right)\right)\right] \\ &= \frac{1}{\sqrt{\lambda_T}} K\left(\frac{x - X_i}{\lambda_T}\right) - \int \frac{1}{\sqrt{\lambda_T}} K\left(\frac{x - u}{\lambda_T}\right) f_X(u)\, du. \end{aligned}$$

Obviously $\mathrm{E}\left(X_{T,k}\right) = 0 \; \forall \; T \in \mathbb{N}, k \in \mathbb{Z}$, but unfortunately $X_{T,i}$ does not fulfill the conditions (7.5.3) and (7.5.4) as it is stipulated in the CLT above. But we can show, that the following two modified conditions are fulfilled instead:

- There exists a constant $C > 0$, such that for all measurable and quadratic integrable functions $f : \mathbb{R}^u \to \mathbb{R}$,

$$\begin{aligned} &\left|\mathrm{Cov}\left(f\left(X_{T,t_1}, \ldots, X_{T,t_u}\right), X_{T,s_1}\right)\right| \\ &\qquad \leq \sqrt{\mathrm{E}\left(f^2\left(X_{T,t_1}, \ldots, X_{T,t_u}\right)\right)} C \lambda_T^{-3/2} \rho_X^r. \end{aligned} \tag{7.5.3'}$$

- There exists a constant $C > 0$, such that for all measurable and bounded functions $f : I\!R^u \to I\!R$,

$$\left|\mathrm{Cov}\left(f\left(X_{T,t_1}, \ldots, X_{T,t_u}\right), X_{T,s_1}X_{T,s_2}\right)\right| \le \|f\|_\infty \, C\lambda_T^{-2}\rho_X^r. \quad (7.5.4')$$

Let us first show that there exists a constant $C \in I\!R$, such that $\mathrm{E}\left(X_{T,k}^2\right) \le C$ and that the conditions (7.5.1), (7.5.2), (7.5.3') and (7.5.4') are fulfilled and sketch afterwards how the proof of Neumann's and Paparoditis' CLT has to be modified to cope with the new situation.

$$\begin{aligned}
X_{T,k}^2 &= \frac{1}{\lambda_T}K^2\left(\frac{x-X_k}{\lambda_T}\right) + \left(\int \frac{1}{\sqrt{\lambda_T}}K\left(\frac{x-u}{\lambda_T}\right) f_X(u)\, du\right)^2 \\
&\quad -\frac{2}{\lambda_T}K\left(\frac{x-X_k}{\lambda_T}\right)\int K\left(\frac{x-u}{\lambda_T}\right) f_X(u)\, du \\
&\le \frac{1}{\lambda_T}K^2\left(\frac{x-X_k}{\lambda_T}\right) + \left(\int \sqrt{\lambda_T}K(s) f_X(x-\lambda_T s)\, ds\right)^2 \\
&\le \frac{1}{\lambda_T}K^2\left(\frac{x-X_k}{\lambda_T}\right) + \lambda_T \|f_X\|_\infty^2 .
\end{aligned}$$

Thus

$$\begin{aligned}
\mathrm{E}\left(X_{T,k}^2\right) &\le \mathrm{E}\left(\frac{1}{\lambda_T}K^2\left(\frac{x-X_k}{\lambda_T}\right)\right) + \lambda_T \|f_X\|_\infty^2 \\
&\le \int K^2(s) f_X(x-\lambda_T s)\, ds + \lambda_T \|f_X\|_\infty^2 \\
&\le \|f_X\|_\infty \|K\|_{L_2}^2 + \lambda_T \|f_X\|_\infty^2 \le C,
\end{aligned}$$

C appropriate, since $\lambda_T \stackrel{T\to\infty}{\longrightarrow} 0$.

Let us now check condition (7.5.1):

$$\left|X_{T,k}\right| / \sqrt{T} \le \epsilon \Leftrightarrow \left|X_{T,k}\right| \le \epsilon\sqrt{T} \Leftrightarrow X_{T,k}^2 \le \epsilon^2 T.$$

Since $X_{T,k}^2 \le \frac{1}{\lambda_T}K^2\left(\frac{x-X_k}{\lambda_T}\right) + \lambda_T \|f_X\|_\infty^2$, there exists a $T_0 \in I\!N$, such that this inequality is fulfilled for all $T \ge T_0$, since $T\lambda_T \stackrel{T\to\infty}{\longrightarrow} \infty$.

Let us now check condition (7.5.2). To do this we will need the following two inequalities (C.2) and (C.3).

On the one hand, we have for $T \in I\!N$, $k \in I\!N$:

$$\begin{aligned}
\mathrm{E}\left(X_{T,0}X_{T,k}\right) &= \mathrm{Cov}\left(X_{T,0}, X_{T,k}\right) = \mathrm{Cov}\left(\frac{1}{\sqrt{T}}K\left(\frac{x-X_0}{\lambda_T}\right), \frac{1}{\sqrt{T}}K\left(\frac{x-X_k}{\lambda_T}\right)\right) \\
&= \mathrm{Cov}\left(g_T(X_0), g_T(X_k)\right) \leq O(1)\mathrm{Lip}\left(g_T\right)\rho_X^k\sqrt{\mathrm{E}\left(g(X_0)^2\right)} \\
&\leq O(1)\lambda_T^{-3/2}\left\|K'\right\|_\infty \rho_X^k \sqrt{\mathrm{E}\left(\frac{1}{\lambda_T}K^2\left(\frac{x-X_0}{\lambda_T}\right)\right)} \\
&\leq C\lambda_T^{-3/2}\left\|K'\right\|_\infty \rho_X^k \sqrt{\|f\|_\infty \|K\|_{L_2}^2} \\
&= O(1)\lambda_T^{-3/2}\rho_X^k.
\end{aligned}$$

Since $(X_{T,k})_{k\in \mathbb{Z}}$ is stationary, we get for $k \in I\!N$:

$$\mathrm{E}\left(X_{T,0}X_{T,-k}\right) = \mathrm{Cov}\left(X_{T,-k}, X_{T,0}\right) = \mathrm{Cov}\left(X_{T,0}, X_{T,k}\right) \leq O(1)\lambda_T^{-3/2}\rho_X^k$$

and so the following inequality is fulfilled for all $0 \neq k \in \mathbb{Z}$:

$$\mathrm{E}\left(X_{T,0}X_{T,k}\right) \leq O(1)\lambda_T^{-3/2}\rho_X^{|k|}. \tag{C.2}$$

On the other hand, we have for all $T \in I\!N$ and all $j, k \in \mathbb{Z}, j \neq k$:

$$\begin{aligned}
\mathrm{E}\left(X_{T,j}X_{T,k}\right) &\leq \mathrm{E}\left(g_T(X_j)g_T(X_k)\right) \\
&= \mathrm{E}\left(\frac{1}{\lambda_T}K\left(\frac{x-X_j}{\lambda_T}\right)K\left(\frac{x-X_k}{\lambda_T}\right)\right) \\
&= \int\int \lambda_T K(u)K(v)f_{X_j,X_k}(u-\lambda_T x, v-\lambda_T x)\,du\,dv \\
&\leq \lambda_T \|K\|_\infty^2 .
\end{aligned} \tag{C.3}$$

Thus for $N(T) \in I\!N$ we get

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} \mathrm{E}\left(X_{T,0}X_{T,k}\right) &= \mathrm{E}\left(X_{T,0}^2\right) + \sum_{\substack{0\neq k\in\mathbb{Z}\\ |k|\leq N(T)}} \mathrm{E}\left(X_{T,0}X_{T,k}\right) + \sum_{\substack{0\neq k\in\mathbb{Z}\\ |k|> N(T)}} \mathrm{E}\left(X_{T,0}X_{T,k}\right) \\
&\leq \mathrm{E}\left(X_{T,0}^2\right) + 2N(T)O(1)\lambda_T + 2\sum_{k=N(T)+1}^{\infty} \lambda_T^{-3/2}O(1)\rho_X^k \\
&= \mathrm{E}\left(X_{T,0}^2\right) + 2N(T)O(1)\lambda_T + \rho_X^{N(T)+1}\lambda_T^{-3/2}O(1)\frac{1}{1-\rho_X}.
\end{aligned}$$

Let $0 < \epsilon < \delta$ and $N(T) := T^{\delta-\epsilon}$, so $N(T)\lambda_T \overset{T\to\infty}{\longrightarrow} 0$ and

$$\rho_X^{N(T)}\lambda_T^{-3/2} = \exp(\ln(\rho_X)\cdot T^{\delta-\epsilon} + \ln(T)(3/2)\delta) \overset{T\to\infty}{\longrightarrow} 0,$$

since $\ln(\rho_X) < 0$. Thus we get

$$\left| \sum_{k \in \mathbb{Z}} \mathrm{E}\left(X_{T,0} X_{T,k}\right) - \mathrm{E}\left(X_{T,0}^2\right) \right| \stackrel{T\to\infty}{\longrightarrow} 0.$$

Furthermore

$$\begin{aligned}
\mathrm{E}\left(X_{T,0}^2\right) &= \mathrm{Var}\left(\frac{1}{\sqrt{\lambda_T}} K\left(\frac{x - X_0}{\lambda_T}\right)\right) \\
&= \int \frac{1}{\lambda_T} K^2 \left(\frac{x-u}{\lambda_T} f_X(u)\, du\right) - \left[\int \frac{1}{\lambda_T} K\left(\frac{x-u}{\lambda_T}\right) f_X(u)\, du\right]^2 \\
&= \int K^2(s) f_X(x - \lambda_T s)\, ds - \left[\sqrt{\lambda_T} \int K(s) f_X(x - \lambda_T s)\, ds\right]^2 \\
&= \int K^2(s) \left(f_X(x) - \lambda_T s f_X'(\theta)\right)\, ds - \lambda_T \left[\int K(s) \left(f_X(x) - \lambda_T s f_X'(\theta)\right)\, ds\right]^2 \\
&= f_X(x) \left\|K\right\|_{L_2}^2 - \lambda_T \int K^2(s) s f_X'(\theta)\, ds - \lambda \left[f_X(x) - \lambda_T \int K(s) s f_X'(\theta)\, ds\right]^2 \\
&= f_X(x) \left\|K\right\|_{L_2}^2 + C_{T,x},
\end{aligned}$$

with

$$|C_{T,x}| \le \lambda_T \left\|K\right\|_\infty \left\|f_X'\right\|_\infty \mathrm{E}\left(|U|\right) + \lambda_T f_X(x) + \lambda_T^2 \left\|f_X'\right\|_\infty \mathrm{E}\left(|U|\right) \stackrel{T\to\infty}{\longrightarrow} 0.$$

Thus

$$\sum_{k \in \mathbb{Z}} \mathrm{E}\left(X_{T,0} X_{T,k}\right) \stackrel{T\to\infty}{\longrightarrow} f_X(x) \left\|K\right\|_{L_2}^2 \tag{C.4}$$

Let $f : \mathbb{R}^u \to \mathbb{R}$ be measurable and quadratic integrable and define $g : \mathbb{R}^u \to \mathbb{R}$, by

$$g(x_1, \ldots, x_u) := f\Big(g_T(x_1) - \mathrm{E}\left(g_T(X_1)\right), \ldots, g_T(x_1) - \mathrm{E}\left(g_T(X_1)\right)\Big),$$

and $h : \mathbb{R} \to \mathbb{R}$ by

$$h(x) := g_T(x) - \mathrm{E}\left(g_T(X_1)\right).$$

Obviously $\mathrm{Lip}\left(h\right) \le \frac{1}{\lambda_T^{3/2}} \left\|K'\right\|_\infty$. Thus

$$\begin{aligned}
\mathrm{Cov}\left(f\left(X_{T,t_1}, \ldots, X_{T,t_u}\right), X_{T,s_1}\right) &= \mathrm{Cov}\left(g(X_{t_1}, \ldots, X_{t_u}), h(X_{s_1})\right) \\
&\le C_1 \sqrt{\mathrm{E}\left(g(X_{t_1}, \ldots, X_{t_u})^2\right)} \mathrm{Lip}\left(h\right) \rho_X^{s_1 - t_u} \\
&= \sqrt{\mathrm{E}\left(f\left(X_{T,t_1}, \ldots, X_{T,t_u}\right)^2\right)} C_1 \frac{1}{\lambda_T^{3/2}} \left\|K'\right\|_\infty \rho_X^{s_1 - t_u} \\
&= \sqrt{\mathrm{E}\left(f\left(X_{T,t_1}, \ldots, X_{T,t_u}\right)^2\right)} C_2 \rho_X^{s_1 - t_u},
\end{aligned}$$

$C_1$ and $C_2$ appropriate. Thus (7.5.3') is fulfilled.
Let now $f : I\!R^u \to I\!R$ be measurable and bounded and define $g : I\!R^u \to I\!R$ like above, by

$$g(x_1, \ldots, x_u) := f\Big(g_T(x_1) - \mathrm{E}\left(g_T(X_1)\right), \ldots, g_T(x_1) - \mathrm{E}\left(g_T(X_1)\right)\Big),$$

and $h : I\!R^2 \to I\!R$ by

$$h(y, z) = \left(g_T(y) - \mathrm{E}\left(g_T(X_1)\right)\right)\left(g_T(z) - \mathrm{E}\left(g_T(X_1)\right)\right).$$

Then obviously $\mathrm{Lip}\,(h) \le (2/\lambda_T^2)\,\|K\|_\infty\,\|K'\|_\infty$, and we get

$$\begin{aligned}
\mathrm{Cov}\left(f\left(X_{T,t_1}, \ldots, X_{T,t_u}\right), X_{T,s_1} X_{T,s_2}\right) &= \mathrm{Cov}\left(g(X_{t_1}, \ldots, X_{t_u}), h(X_{s_1}, X_{s_2})\right) \\
&\le C_3 \sqrt{\mathrm{E}\left(g(X_{t_1}, \ldots, X_{t_u})^2\right)}\mathrm{Lip}\,(h)\,\rho_X^{s_1 - t_u} \\
&\le C_3 \frac{2}{\lambda_T^2}\,\|K\|_\infty\,\|K'\|_\infty\,\|f\|_\infty\,\rho_X^{s_1 - t_u} \\
&= \|f\|_\infty\, C_4 \rho_X^{s_1 - t_u},
\end{aligned}$$

$C_3$ and $C_4$ appropriate. Thus (7.5.4') is fulfilled, too.

It is now left to show, how Neumann's and Paparoditis' proof has to be modified: In Neumann's and Paparoditis' proof the conditions (7.5.3) and (7.5.4) are used to show, that $\left|\sum_{k=1}^n \Delta_{n,k}^{(1)}\right|$ tends to zero, when $n$ tends to infinity. Thus we have to show, that $\left|\sum_{k=1}^n \Delta_{n,k}^{(1)}\right| \overset{n\to\infty}{\longrightarrow} 0$, without using (7.5.3) and (7.5.4).
Notice that Neumann's and Paparoditis' $n$ is our $T$. For the sake of simplicity we will keep our notation and change $n$ to $T$:

We divide $\Delta_{T,k}^{(1)}$ into three parts, just like Neumann and Paparoditis do, but replace their $\Delta$, which is chosen large enough, but constant, with

$$\Delta_T := \log T^\nu,$$

with $\nu$ chosen such that $(2\delta + \nu \log \rho_X) < 0$.
Let $T \in I\!N$ and $k \in \{1, \ldots T\}$. Because of (7.5.3') we have

$$\begin{aligned}
\left|\Delta_{T,k}^{(1,1)}\right| &\le \sum_{j=1}^{k-\Delta_T} \sqrt{\mathrm{E}\left(Y_{T,j}^2\right)} O\left(\frac{1}{\sqrt{T}}\right) C \lambda_T^{-3/2} \rho_X^{k-j} \\
&= \sum_{j=\Delta_T}^{k-1} O\left(\frac{1}{T}\right) \lambda_T^{-3/2} \rho_X^j \le O\left(\frac{1}{T}\right) \lambda_T^{-3/2} \rho_X^{\Delta_T} \sum_{j=0}^{\infty} \rho_X^j \\
&= O\left(\frac{1}{T}\right) \cdot \exp\left(-\frac{3}{2}\log(\lambda_T) + \log(\rho_X)\log(\Delta_T)\right)
\end{aligned}$$

$$
\begin{aligned}
&= O\left(\frac{1}{T}\right)\cdot\exp\left(\frac{3}{2}\delta\log(T)+\log(\rho_X)\cdot\nu\log(T)\right)\\
&= O\left(\frac{1}{T}\right)\cdot\exp\left(\left[\frac{3}{2}\delta+\log(\rho_X)\cdot\nu\right]\log T\right).
\end{aligned}
$$

Thus

$$
\left|\sum_{k=1}^{T}\Delta_{T,k}^{(1,1)}\right| \stackrel{T\to\infty}{\longrightarrow} 0,
$$

since $\left[\frac{3}{2}\delta+\log(\rho_X)\cdot\nu\right]<0$.
Similar to (C.3) one can show $\mathrm{E}\left(g_T(X_j)\right)\le\sqrt{\lambda_T}O(1)$, thus we get:

$$
\begin{aligned}
\left|\Delta_{T,k}^{(1,2)}\right| &\le \sum_{j=k-\Delta_T+1}^{k-1}\mathrm{E}\left(|Y_{T,k}Y_{T,j}|\,2\left\|h^{(2)}\right\|_\infty\right)\\
&\le \sum_{j=k-\Delta_T+1}^{k-1}O\left(\frac{1}{T}\right)\mathrm{E}\left(|X_{T,k}X_{T,j}|\right)\\
&\le \sum_{j=k-\Delta_T+1}^{k-1}O\left(\frac{1}{T}\right)\mathrm{E}\left(\left[g_T(X_k)+\mathrm{E}\left(g_T(X_1)\right)\right]\cdot\left[g_T(X_j)+\mathrm{E}\left(g_T(X_1)\right)\right]\right)\\
&= \sum_{j=k-\Delta_T+1}^{k-1}O\left(\frac{1}{T}\right)\left[3\left(\mathrm{E}\left(g_T(X_1)\right)\right)^2+\mathrm{E}\left(g_T(X_k)g_T(X_j)\right)\right]\\
&\stackrel{(C.3)}{=} \sum_{j=k-\Delta_T+1}^{k-1}O\left(\frac{1}{T}\right)\lambda_T\le\Delta_T O\left(\frac{1}{T}\right)\lambda_T.
\end{aligned}
$$

Thus

$$
\left|\sum_{k=1}^{T}\Delta_{T,k}^{(1,2)}\right|\le O(1)\Delta_T\lambda_T \stackrel{T\to\infty}{\longrightarrow} 0.
$$

$\Delta_{T,k}^{(1,3,3)}$ can be treated just like it is described in Neumann's and Paparoditis' proof. Notice, that different from $\Delta_{T,k}^{(1,2)}$ $\Delta_{T,k}^{(1,3)}$ consists just of one summand, so changing $\Delta$ to $\Delta_T$ doesn't affect $\Delta_{T,k}^{(1,3,3)}$ considerably.
The Lindeberg-Condition (7.5.1), and the boundednes of $h^{(2)}$ and $h^{(3)}$ yield, that for arbitrary $\epsilon>0$:

$$
\begin{aligned}
\left|\sum_{k=1}^{T}\Delta_{T,k}^{(1,3,1)}\right| &\le O\left(\frac{1}{T}\right)\sum_{k=1}^{T}\mathrm{E}\left(X_{T,k}^2 1_{\{|X_{T,k}|\ge\epsilon\sqrt{n}\}}\right)\\
&\quad +O\left(\epsilon\right)\sum_{k=1}^{T}\mathrm{E}\left(|Y_{T,k}|\sum_{j=k-d}^{k}|Y_{n,j}|\right)
\end{aligned}
$$

$$
\begin{aligned}
&= o(1) + O(\epsilon) O\left(\frac{1}{T}\right) \sum_{k=1}^{T} \sum_{j=k-\Delta_T}^{k-1} \mathrm{E}\left(|X_{T,k}| \, |X_{T,j}|\right) \\
&\quad + O(\epsilon) O\left(\frac{1}{n}\right) \sum_{k=1}^{T} \mathrm{E}\left(|X_{T,k}|^2\right) \\
&\le o(1) + O(\epsilon) \sum_{j=T-\Delta_T}^{T-1} \left(3\mathrm{E}\left(g_T(X_1)\right)^2 + \mathrm{E}\left(g_T(X_T) g_T(X_j)\right)\right) + O(\epsilon) \\
&= o(1) + O(\epsilon) \lambda_T \Delta_T + O(\epsilon) \\
&= O(\epsilon),
\end{aligned}
$$

so $\left|\sum_{k=1}^{T} \Delta_{T,k}^{(1,3,1)}\right| \overset{T\to\infty}{\longrightarrow} 0.$

From (7.5.4') we get

$$
\left|\Delta_{T,k}^{(1,3,2)}\right| \le O\left(\frac{1}{T}\right) \lambda_T^{-2} \rho_X^{\Delta_T}.
$$

Thus

$$
\begin{aligned}
\left|\sum_{k=1}^{T} \Delta_{T,k}^{(1,3,2)}\right| &\le O(1) \cdot \lambda_T^{-2} \rho_X^{\Delta_T} = O(1) \exp\left(-2\log(T^{-\delta}) + \Delta_T \log(\rho_X)\right) \\
&= O(1) \exp\left(2\delta \log(T) + \nu \log(T) \log(\rho_X)\right) \\
&= O(1) \exp\left(\log(T) \left(2\delta + \nu \log(\rho_X)\right)\right) \\
&\overset{T\to\infty}{\longrightarrow} 0,
\end{aligned}
$$

since $(2\delta + \nu \log(\rho_X)) < 0$.
Therefore $\left|\sum_{k=1}^{T} \Delta_{T,k}^{(1,3)}\right| \overset{T\to\infty}{\longrightarrow} 0$ and thus

$$
\left|\sum_{k=1}^{T} \Delta_{T,k}^{(1)}\right| \le \left|\sum_{k=1}^{T} \Delta_{T,k}^{(1,1)}\right| + \left|\sum_{k=1}^{T} \Delta_{T,k}^{(1,2)}\right| + \left|\sum_{k=1}^{T} \Delta_{T,k}^{(1,3)}\right| \overset{T\to\infty}{\longrightarrow} 0,
$$

which completes the proof of Theorem 7.5.

□

# Bibliography


[1] Aït Sahalia, Y., Mykland, P. A., Zhang, L. (2005). How Often to Sample a Continous-Time Process in the Presence of Market Microstructure Noise, Review of Financial Studies, 18, 351-416.

[2] Amihud, Y. and Mendelsohn, H. (1987). Trading Mechanisms and Stock Returns: An empirical Investigation. Journal of Empirical Finance, 4, 115-158.

[3] Andersen, T. G. and Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models do provide accurate Forecasts. International Economic Review, 39, 885-905.

[4] Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001). The Distribution of realised Stock Return Volatility. Journal of Financial Economics, 61, 43 - 76.

[5] Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001). The Distribution of Exchange Rate Volatility. Journal of Financial Economics, 61, 43 - 76.

[6] Ango Nze, P., Bühlmann, P. and Doukhan, P.(2002). Weak Dependence beyond Mixing and Asymptotics for nonparametric Regression. Ann. Statist, 30, 397-430.

[7] Areal, N. M. P. C. and Taylor, S. J. (2002). The realised Volatility of FTSE-100 Futures Prices. Journal of Futures Markets, 22, 627 - 648.

[8] Bachelier, L. (1900) Théorie de la spéculation, Annales Scientifiques de l' École Normale Superieure, III-17, 21-86.

[9] Barndorf-Nielsen, O. E., Graversen, S. E. and Shephard, N. (2003). Power Variation & stochastic Volatility: A Review and some new Results. Economics Papers, 2003-W19. Economics Group, Nuffield College, University of Oxford.

[10] Barndorf-Nielsen, O. E. and Shephard, N. (2001). Non Gaussian Ornstein-Uhlenbeck-based Models and some of their Use in Estimating Stochastic Volatility Models. Journal of the Royal Statistical Society, B, 63, 167 - 241.

[11] Barndorf-Nielsen, O. E. and Shephard, N. (2002). Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models. Journal of the Royal Statistical Society, B, 64, 253-280.

[12] Barndorf-Nielsen, O. E. and Shephard, N. (2003). Realised Power Variation and Stochastic Volatility. Bernoulli, 9, 243-265.

[13] Blattberg, R. C. and Gonedes, N. J. (1974). Temporal Behaviour in Commodity Futures Markets, Journal of Business, 47, 244-280.

[14] Black, F. (1976). Noise. Journal fo Finance, 41, 529 - 543.

[15] Black, F. (1976) Studies of Stock Price Volatility Changes. Proceedings of the Business and Economic Statistics Section, American Statistical Association, 177 - 181.

[16] Black, F. and Scholes, M. (1972). The Valuation of Options Contracts and a Test of Market Efficiency. Journal of Finance, 27, 399-418.

[17] Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. Journal of Political Economy, 81, 637 - 654.

[18] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. Journal of Econometrics, 31, 307 - 327.

[19] Bollerslev, T., Chou, R. Y. and Kroner, K. F. (1992). ARCH Modelling in Finance: A Review of the Theory and empirical Evidence. J. Econometrics, 52, 5 - 59.

[20] Bollerslev, T., Engle, R. F. and Nelson, D. B. (1986). ARCH Models. In Engle, R. F., McFadden, D. (Eds.) Handbook of Econometrics, Edition 1, Volume 4, Chapter 49, pp. 2959 - 3038, Elsevier.

[21] Bosq, D. (1996). Nonparametric Statistics for Stochastic Processes. Lecture Notes in Statistics, 110, Springer-Verlag, New York.

[22] Brown, B.M. (1971). Martingale Central Limit Theorems. The Annals of Mathematical Statistics 42, 59 - 66.

[23] Brown, R. (1828). A brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. Phil. Mag. 4, 161-173.

[24] Brown, S. J. (1990). Estimating Volatility. In Figlewski, S., Silber, W. and Surahmanyam, M. (Eds.) Financial Options: From Theory to Practice, pp. 516-537. Homewood, IL: Business One-Irwin.

[25] Butucea, C. and Matias, C. (2005). Minimax Estimation of the Noise Level and of the Signal Density in a Semiparametric Deconvolution Model. Bernoulli, 11, 309-340.

[26] Carroll, R.J. and Hall, P. (1988) Optimal Rates of Convergence for deconvoluting a Density. J. Amer. Statist. Assoc., 83, 1184-1186

[27] Comte, F. and Renault, E. (1998). Long Memory in Continous-time stochastic Volatility Models. Mathematical Finance, 8, 291 - 323.

[28] Coulon Prieur, C. and Doukhan, P. (2000). A triangular Central Limit Theorem under a new weak Dependence Condition. Statist. Probab. Lett., 47, 61-68.

[29] Courtault, J.-M., Kabanov, Y., Bru, B. and Crépel, P. (2000). Louis Bachelier. On the Centenary of the Théorie de la Spéculation. Mathematical Finance, 10, 341-353.

[30] Cox, J. and Rubinstein, M. (1985). Options Markets. Prentice Hall Inc., Englewood Cliffs, NJ.

[31] Dacorogna, M., Gencay, R., Muller, U., Olsen, R. B. and Pictet, O. V. (2001). An Introduction to High Frequency Finance. Academic Press.

[32] Devroye, L. (1989). Consistent Deconvolution in Density Estimation. Canadian J. Statist., 17, 235-239.

[33] Diggle, P. and Hall, P. (1993). A Fourier Approach to nonparametric Deconvolution of a Density Estimate. J. Royal Statist. Soc. Ser. B, 55, 523-531.

[34] Doukhan, P. (1994). Mixing: Properties and Examples. Lecture Notes in Statistics, 85, Springer-Verlag, New York.

[35] Doukhan, P. and Louhichi, S. (1999). A new weak Dependence Condition and Application to Moment Inequalities. Stoch. Proc. Appl., 84, 313-342.

[36] Eddington, A. S. (1913). Of a Formula for correcting Statistics for the Effects of a Known Probable Error of Observation. Mon. Not. R. Astrom. Soc., 78, 359-360.

[37] Efromovich, S. (1997). Density Estimation for the Case of Supersmooth Measurment Error. J. Amer. Statist. Assoc., 92, 526-535.

[38] Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Annalen der Physik, 322, 549-560. Reprint (2004). Ann. Phys., 14, Supplement, 182-193.

[39] Einstein, A. (1956). Investigations on the Theory of Brownian Movement. Dover, New York.

[40] van Es, B., Spreij, P. and van Zanten, H. (2005). Nonparametric Volatility Density Estimation for discrete Time Models. Nonparametric Statistics, 17, 237 - 251.

[41] Engle, R. F. (1982). Autoregressive Conditional Heteroskedastic Models with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50, 987 - 1007.

[42] Engle, R. F. (1995). ARCH - Selected Readings. Oxford University Press, New York.

[43] Fama, E. F. (1965). The Behaviour of Stock Market Prices. Journal of Business 38, 34-105.

[44] Fan, J. (1991). On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems. Ann. Statist. 19, 1257-1272.

[45] Fan, J. (1991). Global Behaviour of Deconvolution Kernel Estimates. Statist. Sinica, 1, 541-551.

[46] Fan, J. (1993). Adaptively local one-dimensional Subproblems with Application to a Deconvolution Problem. Ann. Statist., 21, 600-610.

[47] Fan, J. and Truong, Y. K. (1993). Non-parametric Regression with Errors in Variable. Ann. Statist. 21, 1900-1925.

[48] Franke, J., Härdle, W. and Kreiss, J.-P. (2003). Nonparametric Estimation in a Stochastic Volatility Model. In Akritas, M. G. and Politis, D. N. (Eds.), Recent Advances and Trends in Nonparametric Statistics, pp. 303-313. Elsevier.

[49] Franke, J., Kreiss, J.-P. and Mammen, E. (2002). Bootstrap of Kernel Smoothing in Nonlinear Time Series. Bernoulli, 8, 1-37.

[50] Franke, J., Kreiss, J.-P., Mammen, E. and Neumann, M. H. (2002). Properties of the Nonparametric Autoregressive Bootstrap. Journal of Time Series Analysis, 23, 555-585.

[51] Gänssler, P., Stute, W. (1977). Wahrscheinlichkeitstheorie, Springer-Verlag, Berlin,

[52] Ghysels, E., Harvey, A. C. and Renault, E. (1996). Stochastic volatility. In Rao, C. R. and Maddala, G. S. (Eds.), Statistical Methods in Finance, pp. 119 - 191. North Holland, Amsterdam.

[53] Hansen, P. and Lunde, A. (2005). Realized Variance and Market Microstructure Noise. Available at "http://papers.ssrn.com/sol3/papers.cfm?abstract_id=506542".

[54] Harris, L. (1990). Estimation of Stock Variance and serial Covariance from discrete Observations. Journal of Finance, 50, 1175-1198.

[55] Harris, L. (1991). Stock Price Clustering and Discreteness. Review of Financial Studies, 4, 389-415.

[56] Hasbrouck, J. (2007). Empirical Market Microstructure: The Institutions, Economics and Econometrics of Securities Trading. Oxford University Press Inc, USA. (Announced to appear in March 2007)

[57] Heston, S. (1993). A closed Form Solution for Options with Stochastic Volatility with Applications to Bonds and Currency Options. Review of Financial Studies, 6, 327-343.

[58] Horowitz, J. L. (1998). Semiparametric Methods in Econometrics. Lecture Notes in Statistics, 85, Springer-Verlag, New York.

[59] Horowitz, J. L. and Markatou M. (1996). Semiparametric Estimation of Regression Models for Panel Data. The Review of Economic Studies, 63, 145-168.

[60] Hull, J. and White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. Journal of Finance, 42, 281 - 300.

[61] Hull, J. and White, A. (1988). An Analysis of the Bias in Option Pricing caused by a Stochastic Volatility. Advances in Futures and Options Research, 3, 29-61.

[62] Hull, J. and White A. (1996). Hull-White on Derivatives. Risk Publications, London.

[63] Kotlarski, I. (1967). On characterizing the gamma and the normal Distribution. Pacific J. Math., 20, 69-76.

[64] Kreiss, J. P. and Neuhaus, G. (2006). Einführung in die Zeitreihenanalyse. Springer - Verlag, Heidelberg.

[65] Li, T. and Vuong, Q. (1998) Nonparametric Estimation of the Measurement Error Model using Multiple Indicators. J. Mult. Anal., 65, 139-165.

[66] Liu, M. C. and Taylor, R. C. (1990). A Consistent Nonparametric Density Estimator for the Deconvolution Problem. Canadian J. Statist., 17, 427-438.

[67] Maheu, J. M. and McCurdy (2002). Nonlinear Features of realised FX Volatility. Review of Economics and Statistics, 84, 668 - 681.

[68] Mandelbrot, B. (1963). The Variation of certain Speculative Prices. Journal of Business 36, 394-419.

[69] Masry, E. (1991). Multivariate Probability Density Deconvolution for stationary Stochastic Processes. IEEE Transactions on Information Theory, 37, 1105-115.

[70] Masry, E. (1993). Asymptotic Normality for Deconvolution estimators of multivariate Densities of stationary Processes. Stochastic Processes and their Applications, 475, 53 - 74.

[71] Masry, E. (1993). Strong Consistency and Rates for Deconvolution of multivariate Densities of stationary Processes. Journal of Multivariate Analysis, 44, 47 - 68.

[72] Masry, E. (1996). Multivariate Regression Estimation, local polynomial Fitting for Time Series. Stoch. Proc. Appl. 65, 81-101. Correction (1997), 67, 281.

[73] Meddahi, N. (2002). A Theoretical Comparison Between Integrated and Realized Volatility. Journal of Applied Econometrics, 17, 479-508.

[74] Meister, A. (2004). On the Effect of Misspecifying the Error Density in a Deconvolution Problem. The Canadian Journal of Statistics, 32, 439-449.

[75] Meister, A. (2004). Deconvolution Density Estimation with a Testing Procedure for the Error Distribution. Stuttgarter Mathematische Berichte 2004-004. Univ. Stuttgart.

[76] Mendelsohn, J. and Rice, J. (1982). Deconvolution of microfluorometric Histograms with B splines. J. Amer. Statist. Assoc., 77, 748-753.

[77] Nadaraya, E. A. (1964). On Estimating Regression. Theory Probab. Appl., 10, 186 - 190.

[78] Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Pricing. A new Approach. Econometrica, 59, 347-370

[79] Neumann, M. H. (1997). On the Effect of Estimating the Error Density in Nomparametric Deconvolution.

[80] Neumann, M. H. (2006). Deconvolution from Panel Data with unknown Error Distribution. Preprint, Univ. Jena.

[81] Neumann, M. H. and Paparoditis, E. (2006). Goodness-of-fit Tests for Markovian Time Series Models. Manuscript.

[82] Palkowski, F. (2004). Nichtparametrisches Bootstrap in heteroskedastischen Zeitreihen. PhD Thesis, Technical University of Braunschweig. http://www.digibib.tu-bs.de/?docid=00001666

[83] Praetz, P. D. (1972). The Distribution of Share Price Changes. Journal of Business, 45, 49-55.

[84] RiskMetrics - Technical Document, Fourth Edition (1996), Morgan Guaranty Trust Company of New York.

[85] Robinson, P. M. (1983). Nonparametric Estimators for Time Series. J. Time Ser. Analysis, 4, 185 - 207.

[86] Samuelson, P. A. (1973). Mathematics of Speculative Prices. SIAM Review, 15, 1 - 42.

[87] Shephard, N. (1996). Statistical Aspects of ARCH and stochastic Volatility. In Cox, D. R., Hinkley, D. V. and Barndorff-Nielsen, O. E. (Eds.), Time Series Models in Econometrics, Finance and Other Fields, pp. 1-67. Chapman & Hall, London.

[88] Shephard, N. (2005). Stochastic Volatility. Oxford University Press, Oxford.

[89] Silverman, B. W. (1986). Density Estimation for Statistics and Data Analysis. Chapman & Hall, London.

[90] Staudenmayer, J. and Ruppert, D. (2004). Local polynomial Regression and Simulation-Extrapolation. Journal of the Royal Statistical Society, B, 66, Part 1, pp. 17 - 30.

[91] Stefanski, L. A. (1990). Rates of Convergence of some Estimators in a Class of Deconvolution Problems. Statistical and Probability Letters, 9, 229-235.

[92] Stefanski, L. A. and Carroll, R. J. (1990). Deconvoluting Kernel Density Estimators. Statistics, 21. 169-184.

[93] Stefanski, L. A. and Carroll, R. J. (1991). Deconvolution-based Score Tests in Measurement Error Models. Ann. Statist. 19, 249-259.

[94] Straumann, D. (2005) Estimation in Conditionally Heteroscedastic Time Series Models. Lecture Notes in Statistics, 181, Springer-Verlag, New York.

[95] Straumann, D. and Mikosch, D. (2006). Quasi-MLE in heteroscedastic Times Series: a stochastic recurrence Equations Approach. To appear in Annals of Statistics

[96] Taqqu, M. S. (2001). Bachelier and his times: a conversation with Bernard Bru. Finance and Stochastics, 5, 3–32, 2001.

[97] Taylor, S. J. (1982). Financial Returns Modelled by the Product of Two Stochastic Processes - A Study of Daily Sugar Prices, 1961 - 1979. Time Series Analysis: Theory and Practice 1 (O.D. Andersen ed.), North-Holland Publishing Company.

[98] Taylor, S. J. (1986). Modelling Financial Time Series. John Wiley and Sons Ltd., Chichester.

[99] Taylor, S. J. (1994). Modelling Stochastic Volatility: A Review and Comparative Study. Mathematical Finance 4, 183-204.

[100] Taylor, S. J. and Kingsman, B. G. (1979). An Analysis of the Variance and Distribution of Commodity Price-changes, Australian Journal of Management 4, 135-149.

[101] Thiele, T. N. (1880) Sur la compensation de quelques erreurs quasi-systématiques par la méthodes de moindre carrés. Vidensk. Selsk. Skr. 5. Rk., naturvid. og mat. Afd., 12, 381–408.

[102] Wand, M. P. and Jones, M. C. (1995). Kernel Smoothing. Chapman and Hall, London.

[103] Wasserman, L. (2006). All of Nonparametric Statistics. Springer Science + Businesss Media, Inc., New York.

[104] Watson, G. S, (1964) Smooth regression analysis. Sankhyā Ser. A, 26, 101 - 116.

[105] Zhang, C. H. (1990). Fourier Methods for Estimating Mixing Densities and Distributions. Ann. Statist. 18, 806-831.

[106] Zhang, L., Mykland, P. and Aït-Sahalia, Y. (2005). A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data. J. Amer. Statist. Assoc., 100, 1394-1411

[107] Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. Journal of Business & Economic Statistics, 14, 45-52.

[108] Zwanzig, S. (1999). How to adjust Nonparametric Regression Estimators to Models with Errors in the Variables? Theory of Stochastic Processes, 5, 21, 272-277.

[109] Zwanzig, S. (2003). Nonparametric functional Errors-in-variables Models with additional Information. Preprint.

# Abstract

We consider nonparametric stochastic volatility models in discrete time with unknown distribution of the innovations of the return process. As underlying and not observable volatility process we assume a nonparametric autoregressive structure of first order. We are interested in estimators for this autoregression function. The considered models generalise on one hand parametric autoregressive random variance models, which quite successfully have been applied to financial time series, and on the other hand nonparametric stochastic volatility models for which the distribution of the innovations of the returns is assumed to be known. We make use of the well accepted assumption that volatility changes (rather) slowly. In a first model we deal with the extreme situation that at least two observed returns are based on exactly the same volatility, which brings us to a situation comparable to panel data. Under certain assumptions we can estimate the characteristic function of the distribution of the innovations. We need an estimator of this distribution in order to define a so-called deconvolution kernel estimate for the autoregression function of the volatility process. In this technically demanding situation we achieve consistency of our estimator.
In another situation we assume, that we can observe the volatility disturbed by a noise, which converges to zero in probability whith increasing sample size. Here we investigate nonparametric kernel-smoothers as well and achieve the same asymptotic results for our estimator as in the situation in which we are able to observe the volatility directly.
Furthermore we introduce two models, which fulfill these assumptions: one, which is based on inter-day log-returns and where a quantity which can be identified with integrated volatility follows the autoregressive structure and another one, which is based on inter-day returns and where a daily mean volatility follows the autoregressive structure.
Finally we consider a nonparametrich GARCH(1,1)-model and show asymptotic normality of an estimator of the stationary density of a process following this structure.

# Zusammenfassung

In der vorliegenden Arbeit werden stochastische Volatilitätsmodelle in diskreter Zeit betrachtet. Die Verteilung der Innovationen des Renditeprozesses wird als unbekannt vorausgesetzt. Wir nehmen an, daß der zugrunde liegende nicht beobachtbare Volatilitätsprozeß einer nicht parametrischen autoregressiven Struktur erster Ordnung folgt und sind an einem Schätzer für die Regressionsfunktion interessiert. Die betrachteten Modelle verallgemeinern einerseits parametrische autoregressive Modelle mit zufälliger Varianz und andererseits nichtparametrische stochastische Volatilitätsmodelle, in denen man davon ausgeht die Verteilung der Innovationen des Renditeprozesses zu kennen. Wir machen uns die verbreitete Annahme zu Nutze, daß sich die Volatilität vergleichsweise langsam verändert. In einem ersten Modell nehmen wir den Extremfall an, daß wir mindestens zwei Renditen beobachten können, die auf exakt der gleichen Volatilität basieren.
Unter gewissen Annahmen können wir in dieser Situation die charakteristische Funktion der Verteilung der Innovationen schätzen. Mit Hilfe dieser Funktion läßt sich basierend auf einem Dekonvolutionskern ein Schätzer für die Autoregressionsfunktion des Volatilitätsprozesses angeben. In dieser technisch aufwändigen Situation erzielen wir Konsistenz des Schätzers.
In einer zweiten Situation nehmen wir an, daß wir die Volatilität um ein Rauschen gestört beobachten können, das mit wachsendem Stichprobenumfang stochastisch gegen Null konvergiert. Hier betrachten wir ebenfalls nichtparametrische Kernschätzer und erzielen die gleichen asymptotischen Resultate wie in der Situation mit direkt beobachtbarer Volatilität.
Weiterhin werden zwei Modelle vorgestellt, die eben diese Annahmen erfüllen: Eines davon basiert auf innertäglichen Log-Renditen und eine Größe vergleichbar mit der sogenannter "integrated volatility" folgt der nichtparametrischen autoregressiven Struktur. In dem zweiten Modell, das auf zwischentäglichen Renditen basiert, folgt eine mittlere Tagesvolatilität der autoregressiven Struktur.
Schließlich betrachten wir noch ein nichtparametrisches GARCH(1,1)-Modell und zeigen asymptotische Normalität eines Schätzers der stationären Dichte des Prozesses, der dieser Struktur folgt.

# Lebenslauf

| | |
|---|---|
| Name: | Andreas Dürkes |
| Geburtsdatum: | 2. Juni 1976 |
| Geburtsort: | Frankfurt am Main |
| | |
| 1982-1986: | Bürgermeister-Hardt-Schule Kelsterbach |
| 1986-1995: | Gymnasium am Kurfürstlichen Schloß Mainz, Abitur |
| | |
| 1995-1996: | Zivildienst<br>(Altenwohn- und Pflegeheim Haus Weingarten / Kelsterbach) |
| | |
| 1996-2002: | Studium der Mathematik (Nebenfach Volkswirtschaftslehre) an der Johannes Gutenberg-Universität Mainz |
| März 2002: | Diplom in Mathematik |
| | |
| seit April 2002: | Wissenschaftlicher Mitarbeiter am<br>Institut für Mathematische Stochastik der<br>Technischen Universität Braunschweig |